

OXFORD
LINGUISTICS

Parts of a Whole

Distributivity as a Bridge Between
Aspect and Measurement

LUCAS CHAMPOLLION

linguistics

language

semantics

interface

morphology

phonetics

phonology

cognition

grammar

syntax

OXFORD STUDIES IN THEORETICAL LINGUISTICS

morphology

computation

Parts of a Whole

OXFORD STUDIES IN THEORETICAL LINGUISTICS

GENERAL EDITORS: David Adger and Hagit Borer, Queen Mary, University of London

ADVISORY EDITORS: Stephen Anderson, Yale University; Daniel Büring, University of Vienna; Nomi Erteschik-Shir, Ben-Gurion University; Donka Farkas, University of California, Santa Cruz; Angelika Kratzer, University of Massachusetts, Amherst; Andrew Nevins, University College London; Christopher Potts, Stanford University; Barry Schein, University of Southern California; Peter Svenonius, University of Tromsø; Moira Yip, University College London

RECENT TITLES

51 The Syntax of Roots and the Roots of Syntax

edited by Artemis Alexiadou, Hagit Borer, and Florian Schäfer

52 Causation in Grammatical Structures

edited by Bridget Copley and Fabienne Martin

53 Continuations and Natural Language

by Chris Barker and Chung-chieh Shan

54 The Semantics of Evaluativity

by Jessica Rett

55 External Arguments in Transitivity Alternations

by Artemis Alexiadou, Elena Anagnostopoulou, and Florian Schäfer

56 Control and Restructuring

by Thomas Grano

57 The Interaction of Focus, Givenness, and Prosody

A Study of Italian Clause Structure

by Vieri Samek-Lodovici

58 The Morphosyntax of Gender

by Ruth Kramer

59 The Morphosyntax of Imperatives

by Daniela Isac

60 Sentence and Discourse

edited by Jacqueline Guéron

61 Optimality-Theoretic Syntax, Semantics, and Pragmatics

From Uni- to Bidirectional Optimization

edited by Géraldine Legendre, Michael T. Putnam, Henriëtte de Swart, and Erin Zaroukian

62 The Morphosyntax of Transitions

A Case Study in Latin and Other Languages

by Víctor Acedo-Matellán

63 Modality Across Syntactic Categories

edited by Ana Arregui, María Luisa Rivero, and Andrés Salanova

64 The Verbal Domain

edited by Roberta D'Alessandro, Irene Franco, and Ángel J. Gallego

65 Concealed Questions

by Ilaria Frana

66 Parts of a Whole

Distributivity as a Bridge Between Aspect and Measurement

by Lucas Champollion

For a complete list of titles published and in preparation for the series, see pp. 313–314.

Parts of a Whole

*Distributivity as a Bridge Between
Aspect and Measurement*

LUCAS CHAMPOLLION

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Lucas Champollion 2017

The moral rights of the author have been asserted

First Edition published in 2017

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2016953469

ISBN 978-0-19-875512-8 (Hbk.)

ISBN 978-0-19-875513-5 (Pbk.)

Printed in Great Britain by
Clays Ltd, St Ives plc

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Contents

<i>General preface</i>	ix
<i>Acknowledgments</i>	x
<i>Abbreviations</i>	xvi
1 Overview	1
1.1 Introduction	1
1.2 The central metaphor	4
1.3 Overview of things to come	6
1.4 Ways to read this book	8
2 The stage	10
2.1 Introduction	10
2.2 Notational conventions	11
2.3 Mereology	11
2.3.1 Foundations of algebraic semantics	12
2.3.2 Axiomatizations of classical mereology	16
2.3.3 Mereology vs. set theory	18
2.3.4 Algebraic closure, atomicity, and groups	19
2.3.5 Higher-order properties	23
2.4 Ontology	23
2.4.1 Ordinary objects (type e)	25
2.4.2 Kinds (type e)	26
2.4.3 Events (type v)	27
2.4.4 Intervals (type i)	28
2.4.5 Degrees (type d)	29
2.4.6 Numbers (type n)	32
2.5 Functions	32
2.5.1 Thematic roles (type $\langle v, e \rangle$)	33
2.5.2 Trace functions: runtime and location (type $\langle v, i \rangle$)	35
2.5.3 Measure functions (types $\langle e, d \rangle$ and $\langle v, d \rangle$)	37
2.5.4 Unit functions (types $\langle d, n \rangle$ and $\langle i, n \rangle$)	38
2.5.5 The cardinality function (type $\langle e, n \rangle$)	40
2.6 Nouns	42
2.6.1 Singular count nouns	42
2.6.2 Plural count nouns	43
2.6.3 Measure nouns	46

2.6.4	Group nouns	47
2.6.5	Mass nouns	49
2.6.6	Kind-referring readings	49
2.7	Verbs	50
2.7.1	The Neo-Davidsonian position and its alternatives	50
2.7.2	Lexical cumulativity	52
2.8	Noun phrases	54
2.9	Prepositional phrases	57
2.10	The compositional process	58
3	The cast of characters	61
3.1	Introduction	61
3.2	Pseudopartitives	62
3.3	<i>For</i> -adverbials	65
3.4	Adverbial- <i>each</i> constructions	68
4	The theory	70
4.1	Introduction	70
4.2	What is distributivity?	71
4.2.1	Quantificational distributivity	71
4.2.2	Relational distributivity	71
4.2.3	Predicative distributivity	72
4.3	Distributive constructions	75
4.3.1	<i>For</i> -adverbials are distributive constructions	76
4.3.2	Pseudopartitives are distributive constructions	76
4.4	The components of a distributive relation	76
4.5	The constraints on distributive constructions	78
4.5.1	Capturing predicative distributivity	78
4.5.2	Adverbial <i>each</i>	84
4.5.3	Temporal <i>for</i> -adverbials	86
4.5.4	Pseudopartitives	89
4.6	Unifying the constraints	93
4.7	Compositional implementation	96
4.8	Summary	100
5	Minimal parts	101
5.1	Introduction	101
5.2	The minimal-parts problem	101
5.2.1	The minimal-parts problem for verbal predicates	102
5.2.2	The minimal-parts problem for mass nouns	103

5.3	Previous accounts of the minimal-parts problem	104
5.3.1	Dowty (1979)	104
5.3.2	Hinrichs (1985)	105
5.3.3	Moltmann (1989, 1991)	107
5.3.4	Link (1991b)	108
5.3.5	Borik (2006)	109
5.3.6	Summing up	110
5.4	Stratified reference as a solution to the minimal-parts problem	110
5.5	Summary	116
6	Aspect and space	117
6.1	Introduction	117
6.2	Subregions versus strata	121
6.3	Implementations of both approaches	125
6.3.1	Dowty (1979) and my proposal	125
6.3.2	Krifka (1998)	128
6.4	Comparing the approaches	129
6.4.1	The subregion-based approach	131
6.4.2	The strata-based approach	135
6.5	Summary	137
7	Measure functions	139
7.1	Introduction	139
7.2	Constraints on measure functions	140
7.3	Previous accounts	141
7.4	Strata theory and measurement	145
7.5	Summary	153
8	Covert distributivity	155
8.1	Introduction	155
8.2	Atomic distributivity	160
8.3	Nonatomic distributivity	163
8.4	Reformulating the D and Part operators	171
8.5	Previous work on event-based distributivity operators	175
8.6	The scopal behavior of <i>for</i> -adverbials	179
8.7	Summary	191
9	Overt distributivity	193
9.1	Introduction	193
9.2	Overt distributivity across languages	196
9.3	Relating overt and covert distributivity	209

9.4	Some more complicated cases	217
9.4.1	<i>Each</i> as a PP modifier	217
9.4.2	<i>Jewels</i> distributing over a conjunction of verbs	218
9.4.3	<i>Jewels</i> in subject position	220
9.4.4	Reverse DP-internal distributivity	222
9.5	Distributive determiners	224
9.6	Zimmermann's generalization explained	230
9.7	Previous work: Zimmermann (2002b)	233
9.8	Summary	236
10	Collectivity and cumulativity	239
10.1	Introduction	239
10.2	<i>All</i> and maximality	241
10.3	<i>Numerous</i> -type predicates	243
10.4	<i>Gather</i> -type predicates	245
10.5	Why <i>all</i> distinguishes between <i>numerous</i> and <i>gather</i>	250
10.6	Meaning shift effects triggered by <i>all</i>	254
10.7	From <i>all</i> to <i>each</i> : different granularity levels	258
10.8	<i>For</i> and <i>all</i> block cumulative readings	260
10.9	Open problems	265
10.10	Summary	267
11	Conclusion	269
11.1	Chapter-by-chapter summary	269
11.2	Future work	275
	<i>Appendix: Distributivity operators as repair strategies</i>	280
	<i>References</i>	283
	<i>Index of languages</i>	305
	<i>Index of people</i>	306
	<i>Index of topics</i>	310

General preface

The theoretical focus of this series is on the interfaces between subcomponents of the human grammatical system and the closely related area of the interfaces between the different subdisciplines of linguistics. The notion of ‘interface’ has become central in grammatical theory (for instance, in Chomsky’s Minimalist Program) and in linguistic practice: work on the interfaces between syntax and semantics, syntax and morphology, phonology and phonetics, etc. has led to a deeper understanding of particular linguistic phenomena and of the architecture of the linguistic component of the mind/brain.

The series covers interfaces between core components of grammar, including syntax/morphology, syntax/semantics, syntax/phonology, syntax/pragmatics, morphology/phonology, phonology/phonetics, phonetics/speech processing, semantics/pragmatics, and intonation/discourse structure, as well as issues in the way that the systems of grammar involving these interface areas are acquired and deployed in use (including language acquisition, language dysfunction, and language processing). It demonstrates, we hope, that proper understandings of particular linguistic phenomena, languages, language groups, or inter-language variations all require reference to interfaces.

The series is open to work by linguists of all theoretical persuasions and schools of thought. A main requirement is that authors should write so as to be understood by colleagues in related subfields of linguistics and by scholars in cognate disciplines.

Lucas Champollion presents, in this volume, a unified theory of four core concepts that lie at the syntax-semantics interface: telicity, plurality, count/mass distinctions, and distributivity. He argues that one can take a unified perspective on these by thinking of them as involving a set of basic semantic operations, baked into functional categories such as prepositions and determiners, applying across different dimensions of events, individuals, and substances. The core idea is that if a predicate applies to something, it also applies to its parts along some dimension and down to a certain level of granularity. This perspective synthesizes and expands previous research on these different topics and allows the first theoretical unification of telicity, plurality, count/mass distinctions, and distributivity. Champollion formalizes his approach and shows in detail how it provides new answers to a number of classical problems.

David Adger
Hagit Borer

Acknowledgments

This book is a substantially revised and extended version of my dissertation, Champollion (2010b), which introduced the framework of strata theory centered around the notion of stratified reference. It incorporates subsequent work as described here.

The story of this work begins in the summer of 2008. I had been a graduate student at the University of Pennsylvania for four years, and I was just about to finish a summer internship at the Palo Alto Research Center (PARC). Partly due to its proximity to Stanford University and Silicon Valley, PARC was a great place to do research at the intersection of linguistics and computer science. Among other things, I had been trying to hack some notion of aspect into the natural-language semantic pipeline they were running at the time. I enjoyed this task: I could feel like the real computational linguist that part of me has always wanted to be, and still read semantics papers all day long. Then my supervisors, Cleo Condoravdi and Danny Bobrow, asked me if I wanted to move to the West Coast and turn my aspect project into a dissertation there. None of us knew back then what it would be about and how much computer science there would be in it. In the end, there is none in it at all. I'm deeply grateful to Cleo and Danny for the trust and enthusiasm with which they embarked on this project with me.

Cleo was the natural choice as my dissertation advisor, and I haven't regretted that choice a single time. Perhaps most importantly for me, she left me the freedom to take this enterprise in whatever direction I wanted. This is not to say that she was ever uninterested or not fully engaged, despite her many responsibilities at PARC. Cleo kept offering patient encouragement throughout the many unexpected turns this work took, and she always had the right amount of clear advice at the right time. She introduced me to the PARC and Stanford communities and, over the course of numerous hiking trips, also to the beauty of the Bay Area.

I'm deeply grateful to the chair of my dissertation, Aravind Joshi. He has been unerringly supportive as he sponsored my first stay at Penn as an exchange student, and later on as I found my way through graduate school between formal language theory and formal semantics. He selflessly supported my decision to move away from Penn and from his own research agenda.

The other members of my dissertation committee were Maribel Romero and Florian Schwarz. Maribel drew me right into semantics from the very first class I took with her at Penn in 2004. Her classes were the best I ever took. Her approach to semantics might well shape my own work more than I know. Like everyone else, I looked on in amazement as she changed jobs, moved across the ocean to Konstanz, founded a family, became department chair, and kept an eye on my dissertation all at once. I am grateful to her for her generous help. There's no way to sneak a flawed linguistic

argument past her, though I'm sure this book contains many that she has long given up on trying to correct.

Although Florian Schwarz joined Penn and the committee only shortly before the dissertation was completed, he worked his way into the project with remarkable speed and made an impact on it, especially by challenging the way I originally characterized stratified-reference constraints as presuppositions. I thank him for numerous detailed comments and for his enthusiastic approach to the enterprise. It was a great pleasure to have him on the committee.

Tony Kroch served on the dissertation committee for a good year or so. I'm grateful that he did, and I'm glad he stayed around even after he officially left. I thank him for giving me the benefit of the doubt early on, for encouraging me to be bold, and for not thinking that his efforts were lost on me. Tony never made a secret of his opinion when he thought I was headed the wrong way, and given my personality, I think he deserves credit for that.

I have a lot of people to thank for teaching me linguistics, engaging with my ideas, and making sure that writing the dissertation and this book was not a lonely business. It would be impossible to thank everyone, and if you feel unjustly omitted from what follows, you most likely are.

The late Ellen Prince took me seriously when I was still wet behind the ears, and I hear she put in a good word for me when I wanted to join Penn. I regret not having spent more time with her.

My Master's degree advisor, Mitch Marcus, has been supportive throughout my time at Penn, and it is largely thanks to his and Aravind's efforts that I could navigate the gap between linguistics and computer science.

I am grateful to my fellow graduate students at Penn from 2004 to 2010, particularly Aviad Eilam, Keelan Evanini, Eva Florencio Nieto, Jonathan Gress-Wright, Catherine Lai, Lucy Lee, Laia Mayol, Jean-François Mondon, and Josh Tauberer. They plowed through the endless coursework with me, endured my island violations, and never lost patience with my unending hunger for grammaticality judgments. It's been great living with my linguist roommates Ariel Diertani, Jonathan Gress-Wright, Laurel MacKenzie (now my colleague at NYU), Laia Mayol, and Satoshi Nambu at 4400 Spruce St. I especially thank Satoshi for helping me print and submit the dissertation while I was in Europe. I thank Yanyan Sui for being who she is, for giving me a glimpse of Chinese, and for showing me the many ways to open an orange. Jeeyoung Kim, Angela Lee, and her Highness Leslie Williams of Sansom Place helped me get away from linguistics once in a while and made sure I didn't take myself too seriously.

I am grateful to the faculty and administrators at the linguistics department at Penn, particularly to Gene Buckley and Amy Forsyth, for their support and for making it possible for me to spend time at PARC and Stanford. I also thank Penn for supporting me through a Benjamin Franklin fellowship, a Dissertation Completion fellowship, and various travel grants.

Coming to Penn in the first place wasn't easy, and many people prepared and supported me on the way. I thank my teachers, especially my elementary school teacher, Dieter Hellmann (†), and my high school English teacher, Peter Schneider. At the University of Freiburg, Udo Hahn introduced me to computational linguistics, recommended Penn to me, and helped me get there. I'm grateful to the German Academic Exchange Service (DAAD) for financing my first year of studies at Penn. I thank the Studienstiftung des deutschen Volkes for their support.

Moving on to California, I gratefully acknowledge PARC for providing me with a stipend and an office. Many thanks to the linguists who were working at PARC while I was there, especially to Danny Bobrow, Cleo Condoravdi, Ji Fang, Lauri Karttunen, John Maxwell, and Annie Zaenen.

I met Liz Coppock at PARC when she was working on spatial aspect. Through working with her, I was eventually led to the ideas in Chapter 6. I'm grateful to her for her comments on that chapter and for many conversations on semantics. Scott Grimm was also at PARC and at the Stanford linguistics department when I was there, and offered comments and generous help on pseudopartitives (Chapter 7) and other topics. Also at Stanford, I thank the faculty, especially Johan van Benthem, Paul Kiparsky, Beth Levin, and Chris Potts, for giving me opportunities to sit in their classes and to present my work. I'm grateful to Eric Acton, David Clausen, Alex Djalali, Jason Grafmiller, Chigusa Kurumada, Sven Lauer, Jack Tomlinson, and my other friends at the linguistics department for many fun and thoughtful conversations. I thank Penn and Stanford for letting me take part in the exchange scholar program, which was highly beneficial to this work.

I finished the dissertation in 2010 as a postdoc at the university of Tübingen, where Sigrid Beck, Fritz Hamm, and Gerhard Jäger have been fostering a lively research environment with help from their wonderful administrators, Sonja Haas-Gruber and Beate Starke. I am grateful to all of them and to the linguist friends I made there, particularly Nadine Bade, Vera Hohaus, Anna Howell, and Sonja Tiemann. Sveta Krasikova had unfortunately already left Tübingen by the time I arrived, but she was still there often enough to give me valuable comments.

I applied for my current position at NYU with a research program based on the dissertation. I have pursued it since I joined the department in 2012. This book summarizes its current state, and I believe it is a stable foundation for future theory-building. I am grateful to the remarkably talented students here and to my semanticist colleagues, Chris Barker, Philippe Schlenker, and Anna Szabolcsi, for the vibrant research environment they have created and for many conversations involving various aspects of this work. For helping me in many ways, I am also grateful to the staff, particularly Aura Holguin, Mike Kennedy, Teresa Leung, and Eddie Quiles. Among my nonsemanticist colleagues, I owe special thanks to Chris Collins for his encouragement and for his detailed comments on Chapters 8 and 9. Among the students, I am especially grateful to Hanna Muller and Linmin Zhang for helping me run web surveys

that I occasionally used to confirm (or in one case, as described four paragraphs ahead, to disconfirm) assumptions I had made in 2010 about the meaning and acceptability of various sentences. In this connection, I thank Michael Yoshitaka Erlewine and Hadas Kotek for providing the open-source software package *Turktools* to the semantics community (Erlewine & Kotek 2016).

I have benefited from many discussions in graduate seminars about this work that I taught at NYU in 2013 and 2014, and in summer school courses that I taught at ESSLLI 2012 and at the 2015 LSA summer institute. Teaching these courses was incredibly inspiring, and I am very grateful to the students who took part in them. Some of the changes in this book originated in discussions in these courses. I owe special thanks to Jeremy Kuhn, who took part in the second NYU seminar and who subsequently presented his work on the word *all* as Kuhn (2014). Jeremy's influence is reflected and acknowledged throughout Chapter 10.

I published an overview of strata theory as a target article in the open peer review journal *Theoretical Linguistics* (Champollion 2015c). I thank Manfred Krifka for encouraging me to write that article, and I am indebted to him and to Hans-Martin Gärtner for their help as editors. I am grateful to the authors of the responses to that article (Corver 2015, Doetjes 2015, Link 2015, Piñón 2015, Schwarzschild 2015, Syrett 2015). These responses prompted me to introduce a number of refinements to the theory, as described in detail in my reply article (Champollion 2015b). This book has been updated to reflect these refinements.

A precursor of Chapter 4 was published as Champollion (2009). Two handbook articles, Champollion (to appear) and Champollion & Krifka (2016), are based in part on the dissertation. While preparing them, I have drawn primarily on the background material in Chapters 2 and 4. This has led to revisions to the text, some of which I have incorporated back into this book. The presentation in Chapter 7 draws in part on section 4 of Champollion (2015b) and on section 3 of Champollion (2015c).

Chapters 1, 2, 3, 4, 6, and 7 have been only lightly changed, mostly to improve presentation. Chapter 5 has been partly rewritten. I have removed discussion of frequency adverbs that was tangential to its main subject, and of a generalization (labeled the “sufficiently-many events” observation) based on subtle gradient judgments that I have since then tried and failed to reproduce experimentally in collaboration with Hanna Muller and Linmin Zhang. I have also expanded the section describing my account.

Chapter 8 has undergone significant expansion and changes compared with the dissertation, leading to its publication as an article in the open-access journal *Semantics and Pragmatics* (Champollion 2016a), which is reprinted here with slight modifications. A closely related proceedings paper, Champollion (2013), is not included in this book, but its contents are referenced at the appropriate places.

Chapter 9 is based on work I carried out at the University of Tübingen and at NYU after the dissertation was completed. Its main ideas have appeared as a short proceedings paper (Champollion 2012). Just like the previous one, this chapter has

been published in its current form as a *Semantics and Pragmatics* article (Champollion 2016c) and is reprinted here with slight modifications. For their help with the text of Chapters 8 and 9, I am grateful to the journal editors, particularly Kai von Fintel and Kjell Johan Sæbø, and to the journal reviewers, particularly Malte Zimmermann.

Chapter 10 is based on Chapter 9 of the dissertation. It has undergone substantial revisions and expansions, described in the main text. A part of this chapter overlaps with Champollion (2015c); another part has appeared as a short proceedings paper, Champollion (2016b). Significant parts of the chapter are new and not included anywhere else. The original chapter contained an extensive discussion of dependent plurals, most of which has been cut from this book because it is only marginally related to strata theory.

Chapter 11, the conclusion, has been rewritten from scratch and substantially expanded to include a chapter-by-chapter summary of the book. Some of the suggestions for future work have previously appeared in Champollion (2015c).

Over the last eight years, many linguists have taken the time to converse or correspond with me about various aspects of this book. Many thanks to Alan Bale, Justin Bledin, Adrian Brasoveanu, Benjamin Bruening, Dylan Bumford, Seth Cable, Gennaro Chierchia, Ivano Ciardelli, Chris Collins, Ashwini Deo, Carmen Dobrovie-Sorin, Jean Mark Gawron, Justyna Grudzinska, Zsófia Gyarmathy, Robert Henderson, Petra Hendriks, Paul Hovda, Sonia Kasyanenko, Chris Kennedy, Manuel Križ, Manfred Krifka, Daniel Lassiter, Chris LaTerza, Lisa Matthewson, Jon Ander Mendia, Alice ter Meulen, Friederike Moltmann, Oleg Neroslavsky, David Nicolas, Hazel Pearson, Maria Mercedes Piñango, Chris Piñón, Jim Pryor, Adam Przepiórkowski, Floris Roelofsen, Uli Sauerland, Remko Scha (†), Roger Schwarzschild, Chung-chieh Shan, Alexander Williams, Yoad Winter, Kata Wohlmuth, Linmin Zhang, Ziren Zhou, Thomas Ede Zimmermann, and Eytan Zweig. Likewise, I thank audiences at the 2009 workshop on Construction of Meaning at Stanford, at the 2009 CHRONOS conference on tense, aspect, and modality in Paris, at the 2010 Penn Linguistics Colloquium, at the 2011 workshop on distributivity in Stuttgart, at the 2011 Amsterdam Colloquium, at the 2013 conference on Semantics and Linguistic Theory at UC Santa Cruz, at the 2014 Cornell workshop in linguistics and philosophy, at the 2015 North Eastern Linguistic Society conference at Concordia University, at the ZAS Berlin, and at the Universities of Connecticut, Dublin (Trinity College), Frankfurt, Göttingen, Konstanz, Los Angeles (UCLA), Maryland, Munich (Center for Mathematical Philosophy), Paris 7, Potsdam, Toruń, Tübingen, Utrecht, Warsaw, and Yale, for their feedback.

For native speaker judgments, I thank Meike Baumann, Renée Blake, Isaac Bleaman, Heather Burnett, Woojin Chung, Ivano Ciardelli, Chris Collins, Liz Coppock, Masha Esipova, Hana Filip, Stephanie Harves, Daði Haþþór Helgason, Hildur Hrólfssdóttir, Gianina Iordăchioaia, Zack Jagers, Sonia Kasyanenko, Mike Kennedy, Songhee Kim, Sverrir Kristinsson, Jeremy Kuhn, Chigusa Kurumada, Luciana

Meinking Guimarães, Yohei Oseki, Roumyana Pancheva, Adam Przepiórkowski, Leonor Remédio, Floris Roelofsen, Bujar Rushiti, Kjell Johan Sæbø, Todd Snider, and Gunnar Ingi Valdimarsson.

At Oxford University Press, I thank Lisa Eaton, Julia Steer, and Vicki Sunter for their help throughout the process, as well as the series editors, David Adger and Hagit Borer. I am grateful for the detailed reviews of the manuscript by Stefan Hinterwimmer and by an anonymous reviewer. I have benefited from them while preparing this book.

I have also greatly benefited from Vera Hohaus' many knowledgeable comments on Chapters 1 through 7, and from Sarah Barrett's, Anna Howell's, Ian Pickett's, and Todd Snider's careful proofreading of the entire manuscript at various stages. I thank Vera Zu for her help with the bibliography, and Philip Shushurin for preparing the index. I gratefully acknowledge a Grant-in-Aid by the NYU Center for the Humanities in support of the publication of this book. I also thank Scott Collard, Jill Conte, April Hathcock, Monica McCormick, and Michael Stoller at the NYU libraries for assisting me in the publishing process.

For their hospitality and much more over the last eight years as this work took me around the world, I'm grateful to many friends of mine: in Philadelphia, Julia Deák, Aviad Eilam, Catherine Lai, and Michela Tincani; in Austin, Liz Coppock; in Berlin, Uli Sauerland and Kazuko Yatsushiro; in Paris, Catherine Collin and my aunts and uncles from the Champollion family, Elisabeth, Ginou, and Hervé; in Portland, Svitlana Antonyuk-Yudina and Vadym Yudin; in Vienna, Gregor and Monika Rot; in Tübingen, Anke and Frank Tochtermann; in Neuchâtel, Sveta Krasikova and Ventsislav Zhechev; in Amsterdam, Galit and Adar Weidman Sassoon; in Warsaw, Justyna Grudzinska; in Moscow, Nadia Antonova, Petya Volyak, Masha Tsyurupa, Nick Gribakin, Natasha Korotkova, Vanya Kapitonov, and Ania Grashchenkova. At an early stage in the development of this project, Arnim von Stechow provided generous advice and a chance to take a step back and examine my ideas from the ultimate bird's-eye perspective, in his mountain hut high up in the snow-covered Swiss alps.

And most of all: *meine Mutter*, *mon père*, *Tuğbacım*. I don't know how to thank you. But still: thank you.

Abbreviations

CEM	classical extensional mereology
CP	complementizer phrase
DP	determiner phrase
IP	inflection phrase
LF	Logical Form
NP	noun phrase
NU	natural unit function
PP	prepositional phrase
SDR	stratified distributive reference
SMR	stratified measurement reference
SR	stratified reference
SSR	stratified subinterval reference
URR	Unique Role Requirement
VP	verb phrase

Overview

The central claim of this book is that a unified theory of distributivity, aspect, and measurement for natural language is feasible and useful.

1.1 Introduction

I claim that a number of natural language phenomena from the domains of aspect,¹ plurality, cumulativity, distributivity, and measurement, which are currently treated by separate theories, are in fact intimately related. Previous accounts of these phenomena either fail to generalize appropriately, or live on as limiting cases of a system presented here under the name of *strata theory*. This system is not a radical reorientation of the grammar. By subsuming and building on previous characterizations, strata theory retains much of what has been formerly gained, and provides a unified framework in which new correspondences are drawn between existing concepts.

The road to this claim starts with four semantic oppositions which are closely associated with the domains under consideration. These are the telic/atelic opposition, which is central to the study of aspect; the singular/plural opposition and the count/mass opposition, which are central to the study of plurality and measurement; and the collective/distributive opposition, which is central to the study of distributivity. These oppositions can be formally related to one another. This in itself is not a new insight. It has long been known that there are close parallels between the singular/plural and the count/mass opposition (e.g. Link 1983) and, likewise, between the count/mass and the telic/atelic opposition (e.g. Bach 1986). That these formal parallels can be extended to encompass the collective/distributive opposition has not been explicitly discussed as far as I know, but it is not difficult to do so.

The nature of the parallelism between all these oppositions can be described intuitively in terms of boundedness. Singular, telic, and collective predicates are

¹ “Aspect” is used in the literature to refer to many different things. Throughout this book, I use the term to refer to what has been variously called inner aspect, lexical aspect, temporal constitution, actionality, or aktionsart, as opposed to the phenomenon referred to as outer aspect, grammatical aspect, or viewpoint aspect. Broadly speaking, I understand inner aspect as referring to the telic/atelic opposition, and outer aspect as referring to the imperfective/perfective opposition. Outer aspect is not discussed in this book.

delimited or bounded in ways that plural, mass, atelic, and distributive predicates are not. Making formal sense of the parallelism therefore amounts to characterizing the difference between boundedness and unboundedness. How to do this is one of the central questions which strata theory proposes to answer. I call it the *boundedness question*.

Answering the boundedness question amounts to specifying what it means for a predicate to be atelic, distributive, plural, or to have mass reference. It is not obvious that there should be a single property that is shared by all these predicates. As this book shows, however, it is indeed possible to isolate such a property. The identity of this property can be determined by analyzing a number of nominal and verbal constructions which have one thing in common: each is sensitive to one of the semantic oppositions listed above. These constructions are *for*-adverbials, which distinguish atelic from telic predicates, as in (1); pseudopartitives, which distinguish plurals and mass nouns from singular count nouns, as in (2); and adverbial *each*, which distinguishes distributive from collective predicates, as in (3). I refer to them collectively as *distributive constructions*.

- | | | | |
|-----|----|---|---------------------|
| (1) | a. | John ran for five minutes. | <i>atelic</i> |
| | b. | *John ran to the store for five minutes. | * <i>telic</i> |
| (2) | a. | thirty pounds of books | <i>plural</i> |
| | b. | thirty liters of water | <i>mass</i> |
| | c. | *thirty pounds of book | * <i>singular</i> |
| (3) | a. | The boys each walked . | <i>distributive</i> |
| | b. | *The boys each met . | * <i>collective</i> |

These three constructions form the empirical basis of this book. However, they probably represent only a small sample of distributive constructions. For example, true partitives and comparative determiners accept the same classes of nouns and of measure functions as pseudopartitives do (Schwarzschild 2006). For present purposes, it is enough to focus on the three distributive constructions in (1) through (3), firstly, because they cut across the domains of distributivity, aspect, and measurement, and secondly because each of them is regarded as central to its domain in the sense that any theory of it must account for its behavior. More concretely, *for*-adverbials are regarded as the prime diagnostic of atelicity (Verkuyl 1989); *each* is the standard example of a distributive item (Link 1987b); and pseudopartitives are arguably the most prominent place in which natural language shows its sensitivity to formal properties of measurement (Krifka 1998, Schwarzschild 2006).

The novel angle of this book consists in considering the constructions in (1) through (3) as parts of a whole. Previous work has produced separate theories to account for the behavior of each of these constructions and for the phenomena that they exemplify. The resulting theories are often more limited in scope than they could be. For example,

work on distributivity has focused on how best to formalize distributive readings, rather than on extending the notion of distributivity. Likewise, the study of aspect has concentrated entirely on temporal phenomena, and the study of measurement in natural language has focused largely on mass terms, partitives, and comparatives. This development has obscured the view on the common properties of these constructions. However, this problem is not inherent in the approaches encoded in these theories. Once the connection between distributivity, aspect, and measurement is made formally explicit, it is easy to connect many existing theories to each other, and to extend them to domains beyond the ones in which they have traditionally been applied. One can then combine the strengths of each account, and synthesize them to extend their empirical coverage. This is the motivation behind this book.

The presence of distributive constructions in every one of the domains of interest makes it possible to place strata theory on a solid empirical foundation, because these constructions allow us to operationalize the boundedness question. Instead of asking abstractly what it is that atelic and distributive and mass and plural predicates have in common with each other, we can search for the property that the bold constituents in the acceptable examples in (1a), (2a), (2b), and (3a) have in common, to the exclusion of the unacceptable examples in (1b), (2c), and (3b).

In order to express generalizations over distributive constructions, I will deploy a common terminology. As is explained in more detail in Chapter 4, *Share* refers to the constituent whose denotation is distributed over the parts of the referent of the other constituent, which is called the *Key*. For example, (1a) distributes *ran* (Share) over *five minutes* (Key); (2b) distributes *water* (Share) over *thirty liters* (Key); and (3a) distributes *walk* (Share) over *the boys* (Key). I assume that these components are related by certain functions such as the function *runtime* in (1a), the function *volume* in (2b), and the thematic role *agent* in (3a). I use the term *Map* for these functions, since they always map entities (such as events or substances) associated with the Share to entities (such as intervals, degrees, or individuals) associated with the Key. These terms and relationships are illustrated in Table 1.1.

This approach results in new takes on a large and diverse number of linguistic phenomena, which are brought together here for the first time in one and the same theoretical picture.

TABLE 1.1. A bridge from distributivity to aspect and measurement

Construction	Example	Key	Share	Map
Adverbial <i>each</i>	The boys each walked	the boys	walk	agent
<i>For</i> -adverbial	John ran for five minutes	five minutes	John ran	runtime
Pseudopartitive	thirty liters of water	thirty liters	water	volume

The rest of this chapter outlines the intuition behind strata theory (Section 1.2), gives a brief overview of the contents of the rest of the book (Section 1.3), and closes with a set of suggestions regarding the different ways in which readers could navigate through the book (Section 1.4).

1.2 The central metaphor

The guiding idea behind this book is that the constructions illustrated in (1) through (3) exclude bounded predicates through a parametrized constraint which is introduced into distributive constructions through certain words such as *for*, *of*, and *each*. This constraint is formulated in terms of a higher-order property, *stratified reference*. This property requires a predicate that holds of a certain entity or event to also hold of its parts along a certain dimension and down to a certain granularity. Dimension and granularity are understood as parameters which distributive constructions can set to different values.

The dimension parameter specifies the way in which the predicate in question is distributed. Different settings of this parameter allow one and the same predicate to be atelic but not distributive, or vice versa. When the dimension parameter is set to time, stratified reference applies to atelic predicates, as in (1). When it is set to a measure function like weight or volume, stratified reference applies to mass and plural predicates, as in (2). When it is set to a thematic role like agent, stratified reference applies to distributive predicates, as in (3).

The granularity parameter specifies that the parts in question must be either atomic or simply smaller in size than the whole, as measured along the dimension. This parameter accounts for the differences between distributive constructions over discrete (count) domains, such as adverbial-*each* constructions, and those over domains involving continuous dimensions, such as *for*-adverbials and pseudopartitives.

The names *dimension*, *granularity*, and *stratified reference* are derived from a visual metaphor, which I develop here. Let me stress that I use this metaphor only for the purpose of conveying the intuitions behind strata theory. It does not have any formal status, it does not occur in the formulation of the theory, and it is not claimed to have any psychological or cognitive reality—unlike, for example, the diagrams in the cognitive grammar literature (Langacker 1986).

The metaphor is based on the idea that individuals, substances, and events occupy regions in an abstract space. The dimensions of this space include the familiar spatial and temporal dimensions as well as any measure functions and thematic roles that happen to be defined for the entity. (To understand a thematic role as a dimension, we assume that the individuals that correspond to these roles are ordered in an arbitrary but fixed canonical order, such as the alphabetical order given by their first and last names.) An object whose weight is large corresponds to a region with a large extent along the weight dimension. An event whose agent is a plural entity corresponds to

a region with a large extent along the agent dimension, while an event whose agent is singular corresponds to a region which is not extended along the agent dimension at all. A temporally and spatially punctual event whose thematic roles are all singular entities corresponds to a point. A temporally and spatially punctual event that has plural entities as its agent and theme corresponds to an infinitely thin rectangle that is extended along the agent and theme dimensions.

Consider the old intuition that any atelic predicate has the subinterval property (Bennett & Partee 1972). This property says that whenever a predicate holds at an interval *t*, it also holds at every subinterval of *t*, all the way down to instants. Put in event-semantic terms, a predicate like *run* is atelic because we can “zoom in” to any temporal part of a running event to find another running event. We cannot do that with a telic predicate like *run to the store*, because any temporal part of an event of running to the store that does not include the end point (the store) does not itself qualify as running to the store. In the metaphor, the subinterval property translates to the following picture: any event in the denotation of a predicate that has the subinterval property can be divided into infinitely thin layers that run perpendicular to the time dimension and that are also in the denotation of this atelic predicate. This gives rise to the well-known “minimal-parts problem”: strictly speaking, there are no instantaneous running events, for example. If the subinterval property is to have any viable chance, it must therefore be amended so that the event layers are constrained to be thinner than the whole event, but do not have to be infinitely thin. Formally, this effect is achieved by adding a granularity parameter to the subinterval property. I call these layers *strata*. This name is chosen to remind the reader of geological strata, the layers of rock which can be observed in geological formations in places such as the Grand Canyon. A geological stratum can be just a few inches thick (though not infinitely thin) and extend over hundreds of thousands of square miles. This aspect is mirrored in the theory, where strata are constrained to be thin along one dimension, but may be arbitrarily large as measured in any other dimension.

The metaphor I have used to describe the subinterval property involves layers or strata rather than points or pebbles, because the subinterval property does not constrain any dimensions other than time. This feature is not accidental. While the relevant parts of running events must be short, or thin, in the temporal dimension, they may have plural entities as agents or themes, they may be extended in space, and so on. This view leads to a natural generalization. Normally, geological strata are horizontal, but due to geological movement, they can also be oriented along another dimension. For example, they can run vertically. Similarly, I have introduced the concept of temporal strata as resulting from dividing an event along the temporal dimension, but we can also imagine spatial or “agental” strata—subevents that are constrained based on their spatial extent or based on their number of agents. Once this step is taken, the atelic/telic opposition can be related to the collective/distributive opposition in a Neo-Davidsonian setting. Distributive predicates require any event in

their denotation to be divisible into strata that are constrained to have atomic thickness on the dimension of the appropriate thematic role. For example, any plural event in the denotation of a predicate like *smile* or *read a book* must be divisible into strata that have atomic agents and that belong to the denotation of the same predicate. Lexical predicates like *smile* have this property due to world knowledge, and phrasal predicates like *read a book* can acquire it through a modified version of the distributivity operators known from Link (1987b) and Schwarzschild (1996). Collective predicates like *be numerous* do not satisfy stratified reference on the thematic role of their subjects, because their subjects can be plural entities whose parts are not themselves numerous.

1.3 Overview of things to come

This section briefly previews the contents of the remaining chapters of the book. A more extended summary is found in Chapter 11. Section 1.4 offers a set of suggestions regarding the different ways in which readers can navigate through the book.

Chapter 2, “The stage,” presents a distilled picture of the crucial issues in the theoretical background assumptions, and develops the framework on which strata theory is built. This framework is essentially a synthesis of the work by Lønning (1987), Link (1998a), Krifka (1998), Landman (2000), and others. Its mathematical foundation is classical extensional mereology, which is presented and discussed at length. The overview in this chapter is intended as a reference point for future researchers and spells out the relevant background assumptions as explicitly as possible, especially in the case of choice points where the literature has not yet reached consensus on a preferred analysis. Issues discussed in this chapter include the meaning of the plural morpheme, the question whether the meanings of verbs are inherently pluralized, the formal properties of thematic roles, and the compositional process.

Chapter 3, “The cast of characters,” presents the three constructions listed in (1) through (3) (*for*-adverbials, pseudopartitives, and adverbial *each*) by means of some typical examples. Building on the foundations laid out in Chapter 2, this chapter develops a baseline theory for the syntax and semantics of these constructions and their constituents, keeping things symmetric across domains as much as seems reasonable so that the parallels drawn in subsequent chapters are not obscured more than necessary. The chapter discusses various properties of these constructions, and introduces simplified Logical Forms for them that provide a scaffold on which the theory in the rest of the book is built.

Chapter 4, “The theory,” presents stratified reference as an answer to the bound-
edness question. The parallelism between the telic/atelic, collective/distributive, singular/plural, and count/mass oppositions is captured in a unified framework. After a brief overview of the empirical phenomena that have been discussed under the rubric of distributivity, the notion of stratified reference is gradually developed as a

generalized notion of distributivity. It is then used to formulate a single constraint that explains each of the judgments in (1) through (3), and to predict distributive entailments of lexical predicates via meaning postulates.

Chapter 5, “Minimal parts,” is about the minimal-parts problem: some eventualities and substances fail to distribute at very small scales because they have parts that are too small to satisfy certain mass terms and atelic predicates (Dowty 1979). Focusing on atelic predicates modified by *for*-adverbials, the chapter discusses some previous attempts to solve the problem before presenting a novel solution in detail. It is shown that stratified reference not only avoids problems that infinitely small parts cause for proposals based on the subinterval property and related notions, but also makes the right predictions as far as the interaction between the respective predicate and the length of the interval denoted by the complement of *for* is concerned.

Chapter 6, “Aspect and space,” models the relation between temporal aspect (*run for an hour* / **run all the way to the store for an hour*) and spatial aspect (*meander* / **end for a mile*) previously discussed by Gawron (2009). The chapter shows that *for*-adverbials impose analogous conditions on the spatial domain and on the temporal domain, and that an event may satisfy stratified reference with respect to one of the domains without satisfying it with respect to the other one as well. This provides the means to extend the telic/atelic opposition to the spatial domain. The chapter argues in some detail that stratified reference is in this respect empirically superior to an alternative view of telicity based on divisive reference (Krifka 1998).

Chapter 7, “Measure functions,” explains the linguistic relevance of the difference between extensive measure functions like volume and intensive measure functions like temperature, as illustrated by the pseudopartitives *thirty liters of water* vs. **thirty degrees Celsius of water* (Krifka 1998, Schwarzschild 2006). Subsuming these previous accounts, stratified reference correctly predicts the monotonicity constraint: such constructions disallow measure functions that generally return the same value on an entity and on its parts. For example, in order for **thirty degrees Celsius of water* to be acceptable, it would have to describe a water entity whose parts are colder than itself; but there are no such entities. Stratified reference relativizes unboundedness to just one dimension or measure function at a time. This makes it possible to account for examples like *five feet of snow* even though not every part of a five-foot layer of snow is less than five feet high.

Chapter 8, “Covert distributivity,” investigates and formalizes different sources of covert distributivity. Apart from lexical distributivity effects, which are modeled by meaning postulates, phrasal distributivity is captured via two operators: (i) a D operator distributing over atoms only (Link 1987), and (ii) a cover-based Part operator, which can also distribute over nonatomic pluralities under contextual licensing (Schwarzschild 1996). These operators are couched within strata theory and equipped with its granularity and dimension parameters. The granularity parameter captures

the difference between atomic and nonatomic distributivity. The dimension parameter makes an extension to the temporal domain possible, which explains why indefinites in the syntactic scope of *for*-adverbials tend not to covary (*?John found a flea on his dog for a month*, Zucchi & White 2001). The resulting theory surpasses accounts in which nonatomic distributivity is freely available, or not available at all; furthermore, it correctly predicts differences between lexical and phrasal nonatomic distributivity.

Chapter 9, “Overt distributivity,” explains the crosslinguistic semantic differences between distance-distributive items such as English *each* and German *jeweils* by treating them as overt versions of the atomic and the nonatomic distributivity operator respectively. The proposed analysis explains why *jeweils* can distribute over salient occasions and why this is never possible for *each* (Zimmermann 2002b). It also accounts for the fact that distributive determiners can take part in cumulative readings with items outside of their syntactic scope, and for their ability to interact with nondistributive event modifiers (Schein 1993, Kratzer 2000, Champollion 2010a).

Chapter 10, “Collectivity and cumulativity,” develops a theory of the word *all*. In some respects, such as its incompatibility with certain collective predicates like *be numerous* (Dowty 1987, Winter 2001) and its limited ability to take part in cumulative readings (Zweig 2009), this word is similar to distributive items like *every* and *each*. In other respects, such as its compatibility with collective predicates like *meet* and its ability to license dependent plurals, it differs from distributive items. This tension between the distributive and the nondistributive facets of *all* is resolved via stratified reference, which is used to formulate the semantics of *all* in terms of a subgroup distributivity requirement and to formulate meaning postulates for collective predicates like *meet* that give rise to distributive inferences to subgroups.

Chapter 11 concludes the book by summarizing its main insights and results. A detailed chapter-by-chapter summary provides a bird’s-eye view of strata theory and stratified reference. The summary highlights the conceptual and theoretical moves as well as their empirical payoff. It contrasts the property-based perspective on stratified reference introduced in Chapter 4 and developed in Chapters 5 through 7 with the operator-based perspective that is central to Chapters 8 and 9, and it sketches how both perspectives come into play in Chapter 10. The book concludes with a list of open problems and suggestions for further research, including a brief discussion of connections to other frameworks such as cognitive and conceptual semantics.

1.4 Ways to read this book

This book presupposes graduate-level knowledge of theories of formal semantics of natural language, as can be found in various textbooks such as Heim & Kratzer (1998). Although this book is self-contained, readers who are new to mereology and algebraic semantics may find it useful to consult the following handbook articles:

for an introduction to classical extensional mereology and an overview of algebraic semantics, Champollion & Krifka (2016); and for an empirical overview of distributivity along with collectivity and cumulativity, Champollion (to appear). These articles overlap in part with this book, but they go into more depth on certain issues, such as aspectual composition in the case of Champollion & Krifka (2016) and psycholinguistic findings as well as crosslinguistic facts in the case of Champollion (to appear).

Readers who are already familiar with these topics, or who are chiefly interested in the linguistic issues discussed in the book, may want to skip Chapter 2 and come back to it in order to clarify questions that come up as they read further.

Champollion (2015c), a target article, provides a self-contained overview of the theory in this book, and can be read as such, especially when taken together with some amendments to the theory described in the last section of the reply article, Champollion (2015b). The theory in this book has been updated to take these amendments into account.

Everyone unfamiliar with these papers who would like to understand just one or two parts of the book should start by reading Chapters 3 and 4; the chapters following these two are modular. Readers who are particularly interested in just one of the topics covered in this book—aspect, measurement, and distributivity—may find it useful to concentrate on the following parts: for aspect, Chapters 5 and 6, and Section 8.6; for measurement, Chapter 7; and for distributivity, Chapters 8, 9, and 10. The detailed chapter-by-chapter summary in Chapter 11 may be helpful as a way to get a bird's-eye perspective on the theoretical and empirical coverage of the book.

The stage

2.1 Introduction

Since this book touches on a number of different semantic domains, it relies on many background assumptions. This chapter presents and motivates these assumptions. Readers who are familiar with the literature and who want to follow the main narrative of this book are encouraged to skip or skim it. I suppose it is best not to read it from beginning to end, but to use it as a reference. For this purpose, the following chapters include many back-references to specific sections in this one.

The textbooks by Partee, ter Meulen & Wall (1990) and by Landman (1991), and the review article on mereology by Champollion & Krifka (2016), provide introductions to many of the mathematical concepts described here. The theory of the syntax–semantics interface follows standard assumptions of generative grammar and builds on the textbook by Heim & Kratzer (1998). To the extent that it is not already contained in that textbook, the formal framework presented in this chapter mainly builds on Krifka (1998), Link (1998a), Landman (2000), and the papers leading up to them. In most cases, I adopt the standard position when there is one, and I state my choice explicitly when no standard position has evolved yet.

The first half of this chapter focuses on my assumptions regarding the ontological primitives and the relations that hold between them. Section 2.2 explains the notational conventions used throughout this book. Section 2.3 provides a general introduction to the conceptual and mathematical underpinnings of algebraic semantics and mereology, defines the system known as classical extensional mereology, and relates it to set theory. Section 2.4 discusses the primitive objects at the basis of the various ontological domains I assume: individuals and substances, spatial and temporal intervals, events, degrees, and numbers. The structural relationships between these domains are expressed by partial functions such as thematic roles and measure functions, which are the topic of Section 2.5.

The second half of this chapter is devoted to various types of linguistic constituents and modes of composition. Section 2.6 covers topics related to nouns, such as the count/mass opposition, the singular/plural opposition, and the status of measure nouns and group nouns. In Section 2.7, I describe my assumptions regarding the

semantics of verbs, which I analyze as sets of events closed under sum formation. Noun phrases are covered in Section 2.8 and prepositional phrases in Section 2.9. Finally, Section 2.10 describes my assumptions regarding the compositional process.

2.2 Notational conventions

I use the following typing conventions: t for truth values, e for ordinary objects (meaning individuals and substances, Section 2.4.1), v for events (which include states, Section 2.4.3), i for intervals (temporal and spatial entities, Section 2.4.4), d for degrees (Section 2.4.5), and n for numbers (Section 2.4.6). I use the symbols x, y, z, x', y', z' and so on for variables that range over ordinary objects, e, e', e'' for events, t, t', t'' for intervals, n, n', n'' for numbers, and b, b', b'' for entities of any basic type. I use P for predicates of type $\langle e, t \rangle$, V for predicates of type $\langle v, t \rangle$, and f for functions of various types. I use θ and Θ for functions of type $\langle v, e \rangle$. Some variables range over objects of different types; when this is clear from context, I will continue to use the symbols above. For example, the range of θ will also include runtime, a function from events to intervals (see Section 2.5.2).

I assume that at least individuals, substances, events, and intervals are closed under mereological sum formation (Section 2.3.1). Intuitively, this means that these categories include plural entities. The lowercase variables just mentioned should therefore be taken to range over both singular and plural entities. In the literature on plurals, the distinction between singular and plural entities is often indicated by lowercase and uppercase variables. Since almost all the variables in my representations range over potentially plural entities, I do not follow this convention.

For acceptability judgments, I use the * (star) sign as indicating either syntactically or semantically unacceptable sentences. Some authors reserve * for syntactically unacceptable sentences and use another sign, such as # (hash), for semantically unacceptable ones. I do not follow this convention because most of the unacceptable sentences I mention are semantically rather than syntactically unacceptable. I use ? and ?? to indicate that a sentence has borderline status. Occasionally I give judgments for constituents which are not entire sentences. These judgments should be taken as indicating the status of appropriate sentences in which the constituent is used. I sometimes use % to indicate that acceptability varies from speaker to speaker.

2.3 Mereology

This section provides a general introduction into the conceptual and mathematical underpinnings of algebraic semantics and mereology, defines the system known as classical extensional mereology, and relates it to set theory.

2.3.1 *Foundations of algebraic semantics*

In standard Montague semantics, the domain of discourse (the collection of things from which the denotations of words and larger constituents are built) is simply a collection of disjoint nonempty sets. Many model-theoretic accounts follow this tradition and assume no more than a set of individuals and a set containing the truth values; others also add sets representing such entities as events or possible worlds. Although this type of setup is very simple, it is a solid basis for formalizing the semantics of large areas of natural language. The textbook by Heim & Kratzer (1998) illustrates this approach.

In this book, I use the term *algebraic semantics* to refer to a semantics in which the domain of individuals also includes plural and mass entities (Link 1983, 1998a). These entities are assumed to stand in a relation \leq , called *parthood*. This relation is added directly to the logic and model theory, and its properties are described by axioms.

Several extensions of algebraic semantics have been proposed to include other ontological domains such as intervals and events (Hinrichs 1985, Krifka 1989b, 1998, Link 1987a, Landman 2000). I discuss my ontological assumptions concerning these entities in Section 2.4, which focuses on the axiomatic underpinnings of algebraic semantics.

The philosophical theory of parthood relations is called *mereology*. Both in the philosophical and in the semantic literature, there is no consensus on which general principles should be taken to constrain the meaning of expressions like $a \leq b$, much less what they are supposed to capture conceptually. In this book, I take the position that the relation \leq is no more than a formal device in a semanticist's theory-building toolbox, a part of the metalanguage which should not be seen as subject to any constraints other than the ones the theory explicitly places on it. To help build intuitions, Table 2.1 gives a few examples of what I intend the relation \leq to express.

The examples in Table 2.1 have in common that there is a sense in which the “wholes” are collections without internal structure; they consist exactly of a bunch of things in the category of “parts,” and their properties can be induced from these parts.

TABLE 2.1. Some examples of wholes and parts

Whole	Part
some horses	a subset of them
a quantity of water	a portion of it
John, Mary, and Bill	John
some jumping events	a subset of them
a running event from A to B	its part from A halfway towards B
a temporal interval	its initial half
a spatial interval	its northern half

In mereology, this concept is called *sum*. I call this parthood relation *unstructured parthood*, as opposed to the relation that holds between singular entities and their parts, which I call *structured parthood*.

I use \leq to model only unstructured parthood. I believe that this decision is in line with the algebraic semantic literature. Whether or not natural language uses the word *part* for some relation, I do not consider this to be evidence for the nature of \leq . The relation \leq should not be constrained by whatever is regarded as the meaning of the natural-language expression “part of,” although it is obviously inspired by it. Not all semanticists agree on this point; see for example Moltmann (1997, 1998) for an opposing view, and Pianesi (2002) and Varzi (2006) for responses. For discussion on related philosophical issues, see Simons (1987), Varzi (2016), Champollion & Krifka (2016) and references therein.²

To make sure that the formal relation \leq reflects the conceptual properties of the relevant notion of parthood, a number of axioms are imposed on it. Mereology does not have a standard axiom system. It is therefore important for any work that relies on mereology to specify which mereology is meant.

I adopt *classical extensional mereology* (CEM), a system which is probably the most widely used mereological system in philosophy and linguistics, although the linguistic literature is not always explicit about this. I now present CEM and lay out some of the intuitions behind the use of mereology for semantics. I will define CEM by taking parthood as a primitive relation and formulating axioms that impose constraints on it. Alternatively, we could also start out with the sum operation, and derive from that the parthood relation. Both approaches are described and contrasted in Champollion & Krifka (2016). The discussion in this section is partly based on the excellent surveys in Simons (1987), Casati & Varzi (1999), and Varzi (2016).

The following axioms constrain parthood to be a partial order:

- (1) **Parthood (primitive relation)**
 $x \leq y$
 (x is part of y.)
- (2) **Axiom of reflexivity**
 $\forall x[x \leq x]$
 (Everything is part of itself.)
- (3) **Axiom of transitivity**
 $\forall x \forall y \forall z[x \leq y \wedge y \leq z \rightarrow x \leq z]$
 (Any part of any part of a thing is itself part of that thing.)

² I thank Alexander Williams and Jim Pryor for discussing this issue with me.

(4) **Axiom of antisymmetry**

$$\forall x \forall y [x \leq y \wedge y \leq x \rightarrow x = y]$$

(Two distinct things cannot both be part of each other.)

These requirements are so basic that they are fulfilled not only by CEM but by almost all mereological theories. Despite this, each of them has been criticized at times. Moltmann (1997, 1998) emphasizes that transitivity does not match the natural language use of the *part of* construction. A similar point can of course be made for reflexivity. For example, (5a) and (5b) are odd, but they should be true if *part of* obeyed transitivity and reflexivity, respectively.

- (5) a. This knob is part of this door, this door is part of the doors in my house,
so this knob is part of the doors in my house.
b. John is part of himself.

These facts do not call the use of CEM into question since, as already argued, we do not want to model the everyday usage of the English *part of* construction. My assumptions are not consistent with the relation expressed in (5) because I assume that the elements in the denotations of singular count nouns like *knob*, *door*, *house* and the referents of proper names like *John* are *atomic*: they do not have any parts aside from themselves (see Section 2.6).

Reflexivity is imposed on the parthood relation mainly for technical convenience. We can define an irreflexive *proper-part* relation by restricting parthood to non-equal pairs:

(6) **Definition: Proper part**

$$x < y \stackrel{\text{def}}{=} x \leq y \wedge x \neq y$$

(A proper part of a thing is a part of it that is distinct from it.)

With the part relation, we can define the notion of sum, also called fusion. Sums formally capture the pretheoretical concept of collection: that which you get when you put several parts together. There are several ways to define sum, and while they are equivalent given the CEM axiom system, they are not logically equivalent (Hovda 2009). Mereologies diverge in the way they formalize the concept of sum. I adopt the classical definition of sum in (8), which is due to Tarski (1929). This definition uses the auxiliary concept of overlap, defined as in (7).

(7) **Definition: Overlap**

$$x \circ y \stackrel{\text{def}}{=} \exists z [z \leq x \wedge z \leq y]$$

(Two things overlap iff they have a part in common.)

(8) **Definition: Sum**

$$\text{sum}(x, P) \stackrel{\text{def}}{=} \forall y [P(y) \rightarrow y \leq x] \wedge \forall z [z \leq x \rightarrow \exists z' [P(z') \wedge z \circ z']]$$

(A sum of a set *P* is a thing that contains everything in *P* and whose parts each overlap with something in *P*.)

The definition in (8) is a metalanguage statement about shorthand expansions; that is, it indicates what formula “ $\text{sum}(x,P)$ ” is a shorthand for. Intuitively, this shorthand stands for “ x is a sum of (the things in) P .” If we formulate our theory in first-order logic, the definition in (8) must be interpreted as an axiom schema, that is, a list of axioms in each of which $P(y)$ is instantiated by an arbitrary first-order predicate that may contain y as a free variable. If we formulate our theory in second-order logic, we can instead understand the definition we pick as quantifying over a predicate variable P , which may be interpreted as ranging over sets, as suggested by the paraphrase in (8). In that case, $P(y)$ should be read as $y \in P$ and so on. The set interpretation makes the theory more powerful because there are only countably many formulas, but uncountably many sets (Pontow & Schubert 2006, Varzi 2016).

The formulation of the definition reflects the intuitive fact that a sum may have other parts than just its immediate components. For example, the sum of (i) the referent of the conjoined term *John and Mary* and (ii) the referent of the proper name *Bill* has more parts than these two individuals: it also has the referent of *John* among its parts, as well as the referent of *Bill and Mary*, and so on.

The following facts can be easily shown to follow from these definitions:

(9) **Fact**

$$\forall x \forall y [x \leq y \rightarrow x \circ y]$$

(Parthood is a special case of overlap.)

(10) **Fact**

$$\forall x [\text{sum}(x, \{x\})]$$

(A singleton set sums up to its only member.)

Different mereologies disagree on what kinds of collections have a sum, and whether it is possible for one and the same collection to have more than one sum. In CEM, sums are unique, therefore two things composed of the same parts are identical. This is expressed by the following axiom:

(11) **Axiom of uniqueness of sums**

$$\forall P [P \neq \emptyset \rightarrow \exists! z \text{sum}(z, P)]$$

(Every nonempty set has a unique sum.)

The binary and generalized sum operators in (12) and (13) give us a way to refer explicitly to the sum of two things, and to the sum of an arbitrary set. In these expressions, $\iota x P(x)$ is defined iff there is exactly one object x such that $P(x)$ is true. When defined, the expression denotes that object.

(12) **Definition: Binary sum**

$$x \oplus y \text{ is defined as } \iota z \text{sum}(z, \{x, y\}).$$

(The sum of two things is the thing which contains both of them and whose parts each overlap with one of them.)

(13) **Definition: Generalized sum**

For any nonempty set P , its sum $\bigoplus P$ is defined as $\iota z \text{ sum}(z, P)$.

(The sum of a set P is the thing which contains every element of P and whose parts each overlap with an element of P .)

For example, we can write the plural individual denoted by the conjoined term *John and Mary* as “ $j \oplus m$,” and to the sum of all water as “ $\bigoplus \text{water}$.” Since we can refer to this sum as *the water*, the \bigoplus operator can be thought of as a somewhat similar to a definite description (see Section 2.8).

The pointwise sum operator (14) allows us to construct the sum of a relation. Its main use in this book is as an auxiliary concept in subsequent definitions.

(14) **Definition: Generalized pointwise sum**

For any nonempty n -place relation R_n , its sum $\bigoplus R_n$ is defined as the tuple $\langle z_1, \dots, z_n \rangle$ such that each z_i is equal to

$$\bigoplus \{x_i \mid \exists x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n [R(x_1, \dots, x_n)]\}.$$

(The sum of a relation R is the pointwise sum of its positions.)

2.3.2 *Axiomatizations of classical mereology*

CEM is completely described by the axioms of reflexivity (2), transitivity (3) and antisymmetry (4) taken together with uniqueness of sums (11). In fact, this setup makes axioms (2) and (4) redundant, because any transitive relation that satisfies axiom (11) is provably reflexive and antisymmetric (see Hovda 2009). A typical model of CEM is shown in Figure 2.1.

In general, models of CEM are essentially isomorphic to complete Boolean algebras with the bottom element removed, or equivalently, to complete semilattices with their bottom element removed. This result goes back to Tarski (1935); see also Pontow & Schubert (2006) for qualifications. A bottom element is a “null thing” that is a part of every other thing; the axioms of CEM do not permit the existence of such a thing, apart from trivial models. In the semantic literature, models of CEM are often

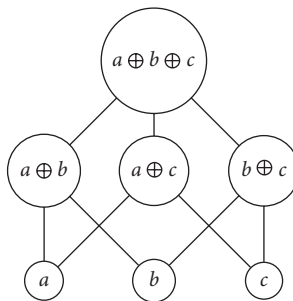


FIG. 2.1 An algebraic structure.

called *complete join semilattices*. However, this use of the term *complete* deviates from standard mathematical practice because of the absence of a bottom element (Landman 1989). I refer to models of CEM as *mereologies*.

(15) **Definition: Mereology**

Let S be a set and \leq be a relation from S to S . A pair $\langle S, \leq \rangle$ is called a mereology iff \leq satisfies the axioms of transitivity (3) and uniqueness (11).

An example of a mereology is the powerset of a given set, with the empty set removed, and with the partial order given by the subset relation. The empty set must be removed because it is a subset of every other set and would therefore act as a bottom element.

The parthood relation described by CEM has essentially the same properties as the subset relation in standard set theory. For practical purposes one can therefore often regard sums as sets, parthood as subsethood, and sum formation as union. Some correspondences between CEM and set theory are listed in Table 2.2. Readers who are unfamiliar with mereology might find this table useful to strengthen their intuitions about the properties of the parthood relation and of the other operations in CEM.

In the Tarski-style axiomatization of CEM adopted here, the sum operation \oplus is defined in terms of the parthood relation \leq , which is taken as a primitive. Properties 2 (transitivity) and 5 (unique sum) in Table 2.2 are considered axioms, and the other properties follow from them as theorems. This is not the only possible way to go. For example, Krifka (1998) uses \oplus instead of \leq as a primitive and imposes the properties 4–9 in Table 2.2 as axioms (he calls 8 the remainder principle). The properties 1–3, which identify \leq as a partial order, then follow as theorems. Other comparable systems are the logic of plurality defined by Link (1983, 1998a), the “part-of” structures of Landman (1989, 1991, 2000), and the lattice sorts and part structures of Krifka (1990, 1998). These systems are clearly intended to describe CEM.

TABLE 2.2. Correspondences between CEM and set theory

Property	CEM	Set theory
1 Reflexivity	$x \leq x$	$x \subseteq x$
2 Transitivity	$x \leq y \wedge y \leq z \rightarrow x \leq z$	$x \subseteq y \wedge y \subseteq z \rightarrow x \subseteq z$
3 Antisymmetry	$x \leq y \wedge y \leq x \rightarrow x = y$	$x \subseteq y \wedge y \subseteq x \rightarrow x = y$
4 Interdefinability	$x \leq y \leftrightarrow x \oplus y = y$	$x \subseteq y \leftrightarrow x \cup y = y$
5 Unique sum/union	$P \neq \emptyset \rightarrow \exists! z \text{ sum}(z, P)$	$\exists! z z = \bigcup P$
6 Associativity	$x \oplus (y \oplus z) = (x \oplus y) \oplus z$	$x \cup (y \cup z) = (x \cup y) \cup z$
7 Commutativity	$x \oplus y = y \oplus x$	$x \cup y = y \cup x$
8 Idempotence	$x \oplus x = x$	$x \cup x = x$
9 Unique separation	$x < y \rightarrow \exists! z [x \oplus z = y \wedge \neg x \circ z]$	$x \subset y \rightarrow \exists! z [z = y - x]$

For example, Link (1983) and Landman (1989) explicitly argue that modeling reference to plurals requires systems with the power of a complete Boolean algebra with the bottom element removed. However, careful review (Hovda 2009) shows that many of these axiomatizations contain errors or ambiguities, mostly in connection with the definition of sum, and therefore fail to characterize CEM as intended. This even extends to the systems given in the standard reference works (Simons 1987, Casati & Varzi 1999). I have tried to avoid these pitfalls by following Hovda's recommendations. Specifically, I follow Hovda's advice in using Tarski's original definition of the sum concept over more modern alternatives.

2.3.3 *Mereology vs. set theory*

Given the correspondences in Table 2.2 and the pitfalls uncovered in Hovda (2009), working directly with set theory instead of mereology might look like a better choice. After all, set theory is better known and more generally accepted than mereology. Indeed, early approaches to plural semantics adopted set theory (Hausser 1974, Bennett 1974). But sticking to mereology has a number of advantages.

First, mereology is used in much of the relevant literature. In the wake of Link (1983), most theories of plurals and mass nouns have been formulated in a mereological framework. This makes it easier to compare and integrate my proposal with existing accounts. (Link's own motivation for adopting mereology over set theory was for philosophical reasons. But Landman (1989), building on Cresswell (1985), argues convincingly that these reasons are independent of linguistic considerations.)

Second, representing plural individuals as proper sums rather than sets is a convenient way to keep them typographically and type-theoretically distinct from the sets that correspond to one-place predicates representing common noun denotations (Vaillette 2001). In this book, both singular and plural individuals are of type *e*.

Finally, moving from mereology to set theory requires some technical adjustments. In order to be able to apply the union operation (which is the counterpart of the sum operation, see Table 2.2) to singular individuals, these individuals need to be replaced by or identified with the singleton sets that contain them (Schwarzschild 1996). In addition, the standard axioms of set theory do not permit infinitely descending chains of set membership of the kind $\dots e''' \in e'' \in e' \in e$. At the bottom of every set membership chain, there must be something that does not have any members itself. In mereological terms, this means that everything is ultimately composed of atoms. A restriction to atomistic mereologies causes problems for the modeling of events, mass entities, and spatiotemporal intervals, where one does not want to be forced to assume the existence of atoms. To avoid this problem, one needs to reject the relevant axioms of set theory. The result is what Bunt (1985) calls ensemble theory.

There is certainly nothing wrong with making these changes to set theory, but taken together, they result in a system that is neither as well known nor as universally

accepted as standard set theory. Depending on the precise axiomatization chosen, its models are likely to turn out to be equivalent to those of CEM. In this sense, we would not gain much by moving from mereology to set theory.

2.3.4 Algebraic closure, atomicity, and groups

Link (1983) proposed algebraic closure as underlying the meaning of the plural:

- (16) a. John is a boy.
 b. Bill is a boy.
 c. \Rightarrow John and Bill are boys.

Algebraic closure (17) extends a predicate P so that whenever it applies to a set of things individually, it also applies to their sum. Algebraic closure was initially introduced in the semantic literature in Link (1983) to describe the meaning of plural formation (see Section 2.6.2) and is indicated by the star operator.

(17) **Definition: Algebraic closure** (Link 1983)

The algebraic closure $*P$ of a set P is defined as $\{x \mid \exists P' \subseteq P[x = \bigoplus P']\}$.

(The algebraic closure of a set P is the set that contains any sum of things taken from P .)

According to this definition, $x \in *P$ means that x consists of one or more parts such that P holds of each of these parts. That is, either P holds of x or else there is a way to divide x into parts such that P holds of each of them. To take an analogy from geometry, if x is a square, and P is the property of being a triangle, then P does not apply to x , but $*P$ does, since any square can be divided into triangles.

Note that the definition in (17) implies that $P' \neq \emptyset$ because \bigoplus is only defined on nonempty sets. Similar provisos apply to the following definitions.

Link translates the argument in (16) as follows:

- (18) $\text{boy}(j) \wedge \text{boy}(b) \Rightarrow * \text{boy}(j \oplus b)$

This argument is valid. Proof: From $\text{boy}(j) \wedge \text{boy}(b)$ it follows that $\{j, b\} \subseteq \text{boy}$. Hence $\exists P' \subseteq \text{boy}[j \oplus b = \bigoplus P']$, from which we have $* \text{boy}(j \oplus b)$ by definition.

The following theorem clarifies the effect of algebraic closure:

(19) **Theorem**

$\forall P[P \subseteq *P]$

(The algebraic closure of a set always contains that set.)

Proof: We need to show that $\forall P[P \subseteq \{x \mid \exists P' \subseteq P[x = \bigoplus P']\}]$, or equivalently, $\forall P \forall x[x \in P \rightarrow \exists P' \subseteq P[\text{sum}(x, P')]]$. This follows for $P' = \{x\}$, given Fact (10).

As the equivalence in (20) illustrates, algebraic closure has a close relation to universal quantification.

(20) **Theorem**

$$x \in {}^*P \Leftrightarrow \forall z[z \leq x \rightarrow \exists z'[P(z') \wedge z' \leq x \wedge z \circ z']]$$

(x is in the algebraic closure of P iff every part of x overlaps with a part of x that is in P .)

The proof of (20) follows from the definitions of sum in (8) and of algebraic closure in (17). Given these definitions, the left-hand side of (20) expands to:

$$(21) \quad \exists C[C \neq \emptyset \wedge C \subseteq P \wedge \forall y[C(y) \rightarrow y \leq x] \\ \wedge \forall z[z \leq x \rightarrow \exists z'[C(z') \wedge z \circ z']]]$$

from which its right-hand side follows directly. For the other direction, we show that $C_x = \{z \mid P(z) \wedge z \leq x\}$ is a witness for C in (21). Clearly $C_x \subseteq P$ and $\forall y[C_x(y) \rightarrow y \leq x]$. Given the right-hand side of (20), reflexivity of \leq entails $C_x \neq \emptyset$, and the definition of C_x entails $\forall z[z \leq x \rightarrow \exists z'[C_x(z') \wedge z \circ z']]$.

Like the sum operation, algebraic closure can be easily extended to predicates of arbitrary arity (Vaillette 2001). The following definition is adapted from the concept of summativity in Krifka (1989b, 1998). I use the symbol \vec{x} to range over sequences. The “double-star operator” ** from the literature on cumulative readings corresponds to the special case for two-place relations. I overload the symbol * and write *R rather than $^{**}R$ because the relation R is not necessarily two-place.

(22) **Definition: Algebraic closure for relations**

The algebraic closure *R of a non-functional relation R is defined as

$$\{\vec{x} \mid \exists R' \subseteq R[\vec{x} = \bigoplus R']\}.$$

(The algebraic closure of a relation R is the relation that contains any sum of tuples each contained in R .)

Finally, the following definition extends algebraic closure to partial functions. While functions are technically relations, it is still useful to treat them separately, because applying the previous definition to a function does not always yield a function. For example, let f_o be the function $\{\langle x, a \rangle, \langle x \oplus y, b \rangle\}$. Applying definition (22) to it yields the relation $\{\langle x, a \rangle, \langle x \oplus y, b \rangle, \langle x \oplus y, a \oplus b \rangle\}$, which is not a function because $x \oplus y$ is related to two distinct values.

The definition in (23) makes sure that the algebraic closure of a partial function is always a partial function. The expression $\lambda x : \varphi . \psi$ represents the partial function that is defined for all x such that φ holds, and that returns ψ wherever the function is defined.

(23) **Definition: Algebraic closure for partial functions**

The algebraic closure *f of a partial function f is defined as

$$\lambda x : x \in {}^*\text{dom}(f). \bigoplus \{y \mid \exists z[z \leq x \wedge y = f(z)]\}.$$

(The algebraic closure of f is the partial function that maps any sum of things each contained in the domain of f to the sum of their values.)

To use the example just mentioned, according to this definition, $*f_o = \{\langle x, a \rangle, \langle x \oplus y, a \oplus b \rangle\}$, which is a partial function.

Other definitions of these concepts exist based on the notion of closure under binary sum formation. For example, an alternative to (17) defines $*P$ as the smallest set P' such that $P \subseteq P'$ and for any two members of P' their sum is also in P' . These definitions are equivalent when P has finite cardinality, but this assumption is not always warranted.³ Take the set $\{x \mid x = x\}$, the set of all things. The axioms of mereology prohibit cycles in the parthood relation. For this reason, if there are only finitely many things, then the domain is *atomic*: every entity is the sum of a set of things which have no proper parts, as defined in (24). However, we do not want to restrict ourselves to atomic domains when we describe spatial and temporal intervals (see Section 2.4.4).

(24) **Definition: Atom**

$$\text{Atom}(x) \stackrel{\text{def}}{=} \neg \exists y[y < x]$$

(An atom is something which has no proper parts.)

By themselves, the axioms of CEM do not specify whether atoms exist or not: among the models they describe, there are some in which everything is made up of atoms (in particular, this includes all models that contain finitely many elements), some in which there are no atoms at all, and intermediate cases. But there are domains for which we do not want to make this assumption. Each of the axioms (25) and (26), though not both at the same time, can be added to CEM to constrain it to one of the two limiting cases. Atomicity entails that everything is ultimately made up of atoms; atomlessness entails that everything is infinitely divisible.

(25) **Optional axiom: Atomicity**

$$\forall y \exists x[x \leq_{\text{Atom}} y]$$

(All things have atomic parts.)

(26) **Optional axiom: Atomlessness**

$$\forall x \exists y[y < x]$$

(All things have proper parts.)

Axiom (25) relies on the following definition, a useful shorthand.

(27) **Definition: Atomic part**

$$x \leq_{\text{Atom}} y \stackrel{\text{def}}{=} x \leq y \wedge \text{Atom}(x)$$

(Being an atomic part means being atomic and being a part.)

Following common practice in algebraic semantics, I assume that the mereology for the count domain is constrained to be atomic. The referents of proper names and the entities in the denotations of singular count nouns are then taken to be absolute

³ I thank Adrian Brasoveanu for reminding me of this point.

mereological atoms. This is in line with my assumptions about the relationship between John and John's arm, or between a ham sandwich and its ham slice. The proper name *John* denotes an atom, and John's arm is not a part of John in the mereological sense. I return to this point in Section 2.6.

In general, however, mereologies do not need to be constrained to be either atomic or atomless. When neither of these axioms is added, the system remains underdetermined with respect to whether or not atoms exist. As Krifka (1998) notes, this underdetermination is one of the advantages of mereology: when we describe the domains of space, time, and mass substances, we do not need to take a stance on such questions as whether atomic events and atomic instants of time exist or whether mass substances can be infinitely subdivided.

One source of terminological confusion is whether atoms should be counted among the sums. On the one hand, it is intuitive to speak of sums only when we consider entities that have more than one part. On the other hand, there is a sense in which atoms are sums: the sum of two atoms x and y can be an atom, namely just in case $x = y$. When I need to avoid ambiguities, I use the term *proper sum* to denote sums that are not atoms. I use the word *entity* to range over atoms and sums.

In Section 2.8, I adopt the notion of *impure atoms* from Landman (1989, 2000) in order to model certain instances of collective predication. Impure atoms, or groups, are atomic entities which are ontologically distinct from sums and derived from them via a special group formation or “upsum” operator, written \uparrow . The operator introduces a distinction between the sum $a \oplus b$, whose proper parts are the individuals a and b , and the impure atom $\uparrow(a \oplus b)$, which has no proper parts. Group formation is a primitive function, that is, it is not defined in terms of other mereological relations. Although it is not strictly speaking part of CEM, I include it here for convenience.

(28) **Group formation (primitive function)**

$$x = \uparrow(y \oplus z)$$

(x is the upsum of $y \oplus z$.)

I call atoms which are not generated through the group formation operator *pure atoms*. The \uparrow operator is constrained such that $\uparrow(a) = a$ iff a is an atomic individual. In other words, atoms are their own groups. The following definitions provide a handle on the pure/impure atom distinction. I assume that impure atoms do not occur in the denotations of singular count nouns—that is, these nouns apply neither to impure atoms nor to entities which have impure atoms as their parts. This assumption has the consequence that even group nouns like *committee* have pure rather than impure atoms in their denotations, as discussed in Section 2.6.4.

(29) **Definition: Impure atom**

$$\text{ImpureAtom}(x) \stackrel{\text{def}}{=} \exists y[y \neq x \wedge x = \uparrow(y)]$$

(An impure atom is an atom that is derived from a distinct entity through the group formation operation \uparrow .)

(30) **Definition: Pure atom**

$\text{PureAtom}(x) \stackrel{\text{def}}{=} \text{Atom}(x) \wedge \neg \text{ImpureAtom}(x)$
 (A pure atom is an atom which is not impure.)

2.3.5 *Higher-order properties*

Algebraic semantics makes essential use of higher-order properties to describe predicate denotations and to model constraints on possible meanings. Various authors have identified the properties of mass nouns with the notions of *cumulative reference* (Quine 1960) and *divisive reference* (Cheng 1973) and the properties of singular count nouns with *quantized reference* (Krifka 1989b), defined below. As described in Section 2.6, I do not rely on the assumption that mass nouns have divisive reference, but I do assume that singular count nouns have quantized reference. Cumulative reference has also been proposed as a property of plural count nouns. It is easy to show that $\text{CUM}(*P)$ holds for any predicate P . I assume that ordinary plural count nouns as well as mass nouns have cumulative reference. Moreover, I assume that verbs, which denote sets of events (Section 2.7), have cumulative reference (Section 2.7.2).

(31) **Definition: Cumulative reference**

$\text{CUM}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[P(y) \rightarrow P(x \oplus y)]]$
 (A predicate P is cumulative iff whenever it holds of two things, it also holds of their sum.)

(32) **Definition: Divisive reference**

$\text{DIV}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$
 (A predicate P is divisive iff whenever it holds of something, it also holds of each of its proper parts.)

(33) **Definition: Quantized reference**

$\text{QUA}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[y < x \rightarrow \neg P(y)]]$
 (A predicate P is quantized iff whenever it holds of something, it does not hold of any its proper parts.)

2.4 **Ontology**

This section is devoted to the relatively extensive list of primitive objects that populate the “ontological zoo,” as we might call it in analogy to the physicists’ particle zoo. I assume that this zoo contains at least individuals and substances, spatial and temporal intervals, events, various sorts of degrees, and numbers. I do not introduce situations or possible worlds, since I do not consider any intensional constructions.

The treatment of the domains of individuals, spatial and temporal intervals, and events is largely symmetric. In particular, these domains are subject to partial orders

which satisfy the axioms of classical extensional mereology. These axioms give rise to algebraic structures like those in Link (1998a), Krifka (1998), and Landman (2000). As Section 2.3 has discussed, they are fairly standard. Degree scales represent a special case, since they are assumed to be totally ordered, while mereologies are typically only partially ordered.

This is a nonreductionist account. I have not tried to reconstruct times from events (Kamp 1979, van Benthem 1983, Landman 1991), events from times (Pianesi & Varzi 1996), events from situations (Kratzer 2016), degrees from individuals (Cresswell 1976), or degrees from numbers (Hellan 1981). With respect to these enterprises, I agree with the sentiment expressed in Link (1983): reductionist considerations are “quite alien to the purpose of logically analyzing the inference structure of natural language . . . Our guide in ontological matters has to be language itself.” A large number of primitives is like a large inventory of tools for theory building: it is easier to work with them, and it is easier to compare the result with other theories that use the same tools. While the framework I adopt here should be in principle amenable to reductionist reconstructions of the sort described, I see them as a subsequent step to linguistic analysis. However, the proliferation of different kinds of entities should not be mistaken for lack of predictive power. Each of these kinds of entities is well motivated through the study of various phenomena. As in the case of the system in Krifka (1998), the predictive power of my system comes mainly from constraining the structural relationships between domains, and not from limiting the number of these domains.

Figure 2.2 illustrates the basic ontology I assume. I use the words *thing* and *entity* to refer to everything in this ontology, abstracting over individuals, events, and so on.

All these domains are assumed to be disjoint from one another. Nothing is both an individual and an event, or both a degree and an interval. This does not preclude a theory that reconstructs some of these entities from each other, as long as some condition similar to disjointness is maintained. Each of the domains consists of one or more sets or sorts, again disjoint. For example, the domain of degrees is subdivided into disjoint sets representing length, weight, temperature, and so on, so that no degree does double duty; and following Link (1983), I assume that count and mass nouns take their denotations in disjoint mereologies. Where exactly to draw the boundaries between one sort and the other is a difficult question. Since I do not rely on the precise nature of these boundaries, I do not try to provide an answer to this question here. Following Krifka (1998) and against Bach (1986), I do not rely on sortal distinctions to model the telic/atelic opposition.

Strictly speaking, each of the mereologies in question comes with its own parthood relation, its own sum operation, and so on. However, since we assume that these mereologies are disjoint, we can easily generalize over these different relations and operations by forming their unions. For example, we form the union of the parthood relation on individuals, the parthood relation on events, and so on, and we let the

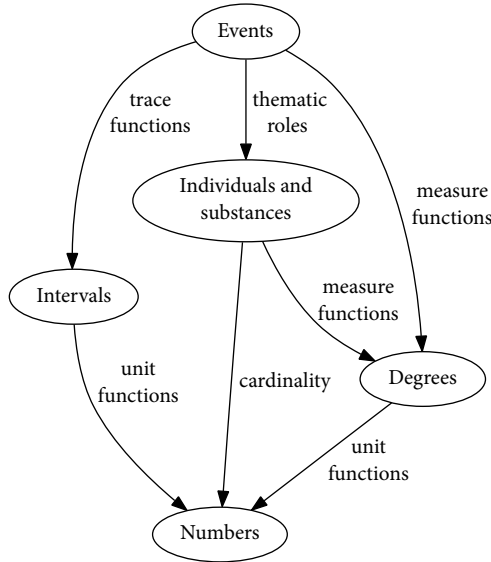


FIG. 2.2 The world (some details omitted).

symbol \leq stand for this union. This is useful to make statements that generalize over mereologies.

I now describe the domains represented in Figure 2.2 in some detail.

2.4.1 Ordinary objects (type e)

The term *ordinary objects* includes anything which can be referred to by a proper name, or denoted by using a common noun, with the exceptions of nominalizations (which arguably involve reference to events) and measure nouns (which are treated here as involving reference to intervals or degrees, see Section 2.6.3). Intuitively, ordinary objects are either individuals or substances. Examples of individuals are firemen, apples, chairs, opinions, and committees (on the topic of group individuals, see Section 2.6.4). Examples of substances are portions of matter such as the water in my cup or the air which we breathe. A basic distinction between individuals and substances is apparently already made by infants, who know that individuals generally have boundaries, move along continuous paths, and survive collisions with each other, while substances do not (Spelke 1990).

Loosely following Quine (1960) and Link (1983), I draw a formal distinction between individuals and substances. The main purpose of this distinction in this book is to express that individuals are built up from mereological atoms, while substances need not be. In connection with the assumption that count nouns always involve reference to individuals, this choice allows us to represent counting and distributing operations as involving the mereological atoms in a plural individual.

This is without doubt a simplifying assumption. For example, so-called fake mass nouns⁴ like *furniture*, *clothing*, *jewelry*, *silverware*, *mail*, and *offspring* provide evidence that at least some mass nouns also involve reference to inherently individuatable and countable entities. Barner & Snedeker (2005) show experimentally that fake mass nouns pattern with count nouns rather than with substance mass nouns in quantity judgment tests. Participants who were asked to compare amounts of matter referred to by either a count noun or an object mass noun relied on number rather than on weight or volume for their judgment. For these reasons, it is difficult to maintain that the referents of fake mass nouns are substances. If fake mass nouns involve reference to atoms as count nouns do, then another factor must be provided that blocks the application of numerals and of distributivity markers to them. I do not provide such a factor in this book. This leaves the potential for overgeneration: I do not have an explanation why one cannot say **three offspring* or **Her offspring each went to a boarding school* (Kratzer 2007). However, let me briefly sketch what such an explanation could look like. An account of fake mass nouns is proposed in Chierchia (1998a, 2010). The atoms involved in the denotation of fake mass nouns are there taken to be “unstable,” that is, they have vague criteria of identity. Numerals are assumed to contain not only an atomicity test but also an anti-vagueness condition that rejects such unstable atoms. By adding the anti-vagueness condition to numerals and *each*, my account could conceivably be extended to prevent such overgeneration. That is, any lexical entries that use the predicate *Atom* would be amended to use a predicate *StableAtom* instead. This predicate would be defined as in Chierchia (2010) and apply only to the individuals in the denotation of count nouns, but not to those in the denotation of fake mass nouns and other mass nouns. However, I do not pursue this route here.

2.4.2 *Kinds (type e)*

Kinds do not play a major role in this book, and their ontological status is debated. Prototypical examples of kinds are plant and animal species, but the inventory of kinds might be much larger; arguably, (almost) all bare plural and mass noun phrases can be understood as involving reference to kinds. While I have categorized kinds as entities in their own right, following Carlson (1977), there have been attempts to formally relate them to other entities. In a mereological setting, it is tempting to associate kinds with sums. For example, the kind *potato* would be the sum of all potatoes. I adopt this idea for convenience; however, I note that it causes problems in the case of kinds such as *Dodo* or *Phlogiston* that are not instantiated in the real world—a problem which can be countered to some extent by appealing to intensionality (Chierchia 1998b), though

⁴ This term comes from Chierchia (1998a, 2010). Other terms include *object mass nouns* (Barner & Snedeker 2005), *collective mass nouns* (Bunt 1985), *count mass nouns* (Doetjes 1997), and *individual mass nouns* (according to Chierchia 2010).

there remain some problems with assimilating kinds to sums too closely (Pearson 2009). I discuss kinds briefly again in Section 2.6.

2.4.3 Events (*type v*)

An event is a spatially and temporally bounded, ephemeral constituent of the world that has a single occurrence (Carlson 1998). The classical examples in the literature of events are Jones' buttering of the toast and Brutus' stabbing of Caesar. As described in what follows, I adopt a standard view according to which sentences existentially quantify over events (Davidson 1967, Parsons 1990), and I assume that events are ordered in a mereological structure (Bach 1986, Krifka 1998). My purpose in making assumptions about events is to bring out the parallel between individuals and events. The null assumption is that there is no semantic distinction between nouns and verbs, only a syntactic distinction.

My notion of events is essentially that of Eckardt (1998), Krifka (1998), and Parsons (1995). In particular, typical events are individual rather than generic, concrete rather than abstract, and located in spacetime.

Just as individuals are "multidimensional" in the sense that they may have several thematic roles and can be measured along several dimensions, events are "multidimensional" in the sense that they exist in space and in time. For example, Caesar's assassination occurred on March 15, 44 BC, in the Roman senate. In this sense, events are distinct from intervals, which are one-dimensional entities and live completely in time or in space (Section 2.4.4). The distinction between events and individuals is more difficult to draw. A popular assumption is that this distinction follows the lines of the 3D/4D controversy (Markosian 2014). Non-instantaneous events have temporal parts (they are 4D objects or perdurants), but all the proper parts of an individual are present at each point in time (they are 3D objects or endurants), or at least throughout the individual's lifetime. For example, the event in which John sleeps from sunset to midnight is a proper part of the event in which John sleeps from sunset to sunrise, but in each of these events, it is John as a whole who sleeps: any and all of his parts are present at every period of time during his sleep, and beyond. This assumption is implicitly present in much work on natural language semantics (e.g. Krifka 1998), and I adopt it too.

I use the term *event* in a wide sense synonymously with *eventuality* (Bach 1986). I assume that all verbs without exceptions involve reference to events (see Section 2.7.1). Other authors use *event* in contradistinction to such things as *states* and sometimes *processes* (Mourelatos 1978, Bach 1986, Smith 1997). While Piñón (1995) argues that states, processes, and events are pairwise disjoint classes, I do not make use of this distinction. I assume, following the arguments in Parsons (1987, 1990: ch. 10), that any declarative sentence, including stative sentences, involves reference to an underlying entity that I call an event. Therefore, the category of events does not only include things like Brutus' stabbing of Caesar, but also things traditionally

called “states,” like John’s being asleep, or the Golden Gate bridge’s spanning from San Francisco to Marin County.

I assume that events form a mereology, meaning that they are closed under sum formation (Bach 1986, Krifka 1998). For example, if John runs to the store, this is an event; if he runs to the store several times, these are several events, and the sum of these events is again an event. Events may have parts which occupy less time, less space, or both. For example, the event e_0 in which John runs from the village to the store has as one of its parts the event e_1 in which John runs from the village halfway to the store. In general, it is not always the case that any part of an event which takes less time also occupies less space. For example, suppose that e_0 is a part of an event e_2 in which John runs from the village to the store and then back. Then e_0 takes less time than e_2 but occupies the same space.

Some authors assume that the domain of events is atomic (Landman 1996, 2000). Others assume that at least some kinds of events have mass denotations, which is especially plausible for events associated with atelicity (Mourelatos 1978). In line with Krifka (1998), I remain neutral on this question, and I do not rely on either assumption.

I assume that different lexical predicates such as singing and tooth-brushing are disjoint, meaning that no event can be both a singing and a tooth-brushing event. An exception involves subtypes of events: a walking event is also a moving event, for example. I assume that certain kinds of event modifiers can provide clues about the identity of events. For example, if I both brush my teeth and sing, but if my singing is loud while my brushing my teeth is not, then the singing is not the same event as the tooth-brushing (Parsons 1990: 157). Not all event modifiers can provide clues about individuation of events. For example, my arriving at my wedding might not be surprising, but my arriving at my wedding late would certainly be; but we would not want to distinguish these as two separate events (Zsófia Gyarmathy, p.c.). Therefore, *surprisingly* is not a suitable choice for the individuation test. Among the event modifiers that are suitable we find manner adverbs like *loudly*, *carefully*, *shyly* and temporal expressions like *from 2pm to 4pm* (Eckardt 1998: chs 1 and 4).

2.4.4 *Intervals (type i)*

For the purpose of this book, intervals are one-dimensional entities that represent stretches of time and paths through space. Examples of intervals are the time from 2pm to 4pm today or the path along which Route 66 extends. I make the standard assumption (as discussed e.g. in Krifka 1998) that intervals are at the basis of the denotations of temporal and spatial measure phrases like *three hours* and *three miles*, and of prepositional phrases like *from 2pm to 4pm* and *to the store* (see Section 2.9). Intervals are also used to specify the spatial and temporal extents of events and other entities. The functions that relate events to intervals are discussed in Section 2.5.2. Intervals also play a separate role in the literature on degree semantics (see Section 2.4.5), but

I set this fact aside here. Also, while one of the main uses of interval semantics is the modeling of tense, I ignore the contribution of tense in this book.

An important distinction between intervals and events is that two distinct events may happen at the same place and/or time, while this is not possible for intervals. For example, a sphere can rotate quickly and heat up slowly at the same time (Davidson 1969). Since the predicates *slowly* and *quickly* exclude each other, I assume that they apply to different events, but then these are distinct events that happen at the same time and at the same place.

The standard conception of time in natural language semantics is based on intervals, which are generally conceived as being ordered by a subinterval relation and a temporal precedence relation. The relevant choices for axiomatizing temporal structures are thoroughly examined and discussed by van Benthem (1983). The algebra of paths and the contribution of directional prepositional phrases to lexical aspect has been studied by Zwarts (2005), among others. A summary of the literature can be found in Ursini (2006). Here I follow Krifka (1998), who proposes axioms for both temporal and spatial interval structures and integrates them in a mereological system, based on earlier ideas by Hinrichs (1985). Krifka's axioms ensure, among other things, that temporal precedence is irreflexive, asymmetric, and transitive, and that any two intervals either overlap or one precedes the other.

The subinterval relation is equated with mereological parthood and written \leq ; I write the temporal precedence relation as \ll . For example, if a is the interval from 2pm to 3pm today, b is the interval from 4pm to 5pm, and c is the interval from 1pm to 5pm, then we have $a \leq c$, $b \leq c$, and $a \ll b$. I assume that there are also directed spatial intervals. For example, the prepositional phrase *to the store* involves reference to a directed interval whose end is the store (see Section 2.9). I use the functions *start* and *end* to refer to appropriate initial and final parts of directed intervals.

Since time is assumed to be a mereology, any two intervals have a unique sum. Since intervals can be arbitrarily summed, some of them are discontinuous. In our example, the interval $a \oplus b$ is discontinuous: it contains a gap from 3pm to 4pm. For this reason, the term "interval" is perhaps misleading, but I use it anyway following common practice.

I assume that the mereologies underlying intervals are nonatomic. According to von Stechow (2009), this is a standard assumption: most semanticists do not assume the existence of temporal atoms or instants.

2.4.5 Degrees (type d)

I assume that the ontological system contains a category of entities called degrees, which represent quantities assigned by measure functions such as height, weight, volume, or beauty. Examples of degrees are John's weight, the thickness of the ice at the South Pole, and the speed at which John is driving his car now. Degrees have been used in semantic accounts of gradable adjectives (*tall*, *beautiful*) and measure

nouns (*liter, hour*) as well as in accounts of the constructions that contain them, such as measure phrases (*three liters*), comparatives (*taller, more beautiful, more water, more than three liters*), and pseudopartitives (*three liters of water*). Degree semantic approaches to comparatives interpret *Mary is taller than John* as some variation of *Mary's height exceeds John's height*, where Mary's height is the degree to which she is tall. Similarly, *Mary is more beautiful than John* is interpreted as *Mary's beauty exceeds John's beauty*, where Mary's beauty is the degree to which she is beautiful. Degree semantic approaches to measure phrases interpret *three liters* as involving reference to a degree of volume. As we will see in Section 3.2, a pseudopartitive like *three liters of water* is then interpreted as involving reference to a water entity whose degree of volume is specified by *three liters*.

In the modern literature, semantic analyses that treat degrees as part of the ontology originate with Cresswell (1976). A different line of analysis treats degrees as contextual coordinates rather than as ontological entities in their own right (Lewis 1972). For comparisons of the two approaches, see Klein (1991), Kennedy (2007), and van Rooij (2008). I do not adopt the contextual-coordinate analysis because it is mainly concerned with gradable adjectives and does not, as far as I can see, provide an obvious way to represent the meaning of measure phrases.

Within analyses that treat degrees as part of the ontology, there is disagreement on what these degrees are. I model degrees as primitive entities. This is also the option chosen by Parsons (1970) and Cartwright (1975). Other authors model degrees as numbers (Hellan 1981, Krifka 1998) or as equivalence classes of individuals indistinguishable with respect to a relevant gradable property (Cresswell 1976, Ojeda 2003). There seems to be a consensus that degrees of a given sort are totally ordered, although it is a topic of debate whether degrees should be understood as points or initial intervals ("extents") on a scale. (Such intervals must be carefully distinguished from the temporal and spatial intervals in Section 2.4.4. Temporal and spatial intervals are ordered by two relations: precedence and parthood. Degree-based intervals are assumed to be initial intervals on their scale and are therefore only ordered by a parthood relation.)

I adopt the view that degrees are totally ordered on scales. Following Fox & Hackl (2006), I assume that this order is dense. I represent it as \leq , just like the mereological parthood relation. I leave open the question whether degrees should be modeled as points or extents, as it does not seem to affect my proposal (see Krasikova 2010). While I generalize over degree order and mereological parthood, I remain noncommittal on the question of whether degree scales should indeed be understood to be special cases of mereologies (see Szabolcsi & Zwarts (1993) and Lassiter (2010a, 2010b) for a proposal in this direction). While there is nothing wrong conceptually with assuming that the ordering relation on degrees of a scale is a special case of the mereological parthood relation, the former relation is usually assumed to be a total order while mereological parthood, at least in the general case, is only constrained to be a partial

order. This has the consequence that the mereological sum operation does not have a natural counterpart on degree scales. Many ontological systems assume that there is a sum operation on degrees, which corresponds to arithmetic sum. This degree sum operation has different properties from mereological sum. For example, mereological sum operation is idempotent (Table 2.2), but this is not the case for the degree sum operation. Imagine two coins which weigh one gram each. The degree sum of their weights is two grams. However, if the two coins are assumed to have identical weight degrees, then the mereological sum of their weights is one gram, because sum is idempotent. For more on this point, see Section 2.5.3.

Let me briefly explain why I do not derive degrees from other entities. Reconstructing degrees as equivalence classes does not help integrating them into a mereological system because it does not lead to any straightforward set-theoretic or mereological relations between different degrees. For example, Cresswell (1976) reconstructs the degrees *three pounds* and *six pounds* as the sets of all things whose weight is three pounds and six pounds, respectively. This means that *three pounds* is neither an element nor a subset of *six pounds*. The operation of adding *three pounds* to *three pounds* cannot be easily interpreted mereologically either, supposing that the result can be described as *six pounds*. It is different from mereological sum formation, since that operation is idempotent. It is also different from pointwise sum formation: although some of the elements of the set *six pounds* are mereological sums of two elements of the set *three pounds*, others might be atoms.

Reconstructing degrees as numbers would raise problems for modeling unit conversion. For example, 6ft approximately equals 183cm, but 6 is not equal to 183. It would also make it difficult to model the difference between different scales such as height or weight, since one set of numbers is used for all purposes. On the other hand, the total ordering and other properties of numbers are useful to model mathematical properties of degrees. A compromise proposed by Lønning (1987), which I adopt, is to introduce degrees as an intermediate layer that mediates between individuals and numbers. In Lønning's framework, one set of functions maps entities to their degrees of weight, length, temperature, and so on. Another set of functions maps these degrees to numbers or number-like entities. With the help of the intermediate layer, we can model degrees of weight, height, temperature and so on as different sorts. See Section 2.5.4 for discussion.

Having both degrees and intervals may seem redundant. Both time and space could be thought of as degree scales. However, I keep temporal and spatial intervals formally apart from degrees because time and space appear to be qualitatively different from degree scales. We can individuate temporal and spatial intervals, but we cannot do anything corresponding with degrees. For example, we feel that Thursday is a different temporal interval than Friday, even though both have the same length. But there is no sense in which we feel that two degrees that represent the same weight or temperature are different from each other. Intervals are subject to two partial orders:

precedence and inclusion. For example, Thursday precedes Friday, and Thursday includes Thursday afternoon. Degrees only appear to be subject to one (total) order: four kilograms is less than five kilograms.

Another difference concerns gaps. As described in Section 2.4.4, I assume that intervals can be discontinuous. I make no such assumption for degrees, however. This makes intuitive sense on the view that identifies degrees with points on a totally ordered scale. Empirically, it is justified by the differing behavior of event descriptions such as *work for three hours* or *three hours of work* on the one hand, and substance descriptions such as *three liters of oil* on the other. An event that qualifies as *three hours of work* may be punctured by interruptions (see Section 2.5.4 for more discussion of this point). By contrast, as a reviewer notes, in order for a substance-based pseudopartitive like *three liters of oil* to apply to a substance, it is not sufficient that the volume of a substance is three liters and that it consists of oil for the most part; it really needs to be pure oil (at least up to the prevailing standard of precision, see Lasersohn 1999).

2.4.6 Numbers (type *n*)

I use numbers to represent the meanings of number words. For example, the word *three* denotes the number 3. Numbers also constitute the range of what I call unit functions, that is, functions like *hours* and *meters* (see Section 2.5.4). I use the type *n* to represent them. I assume that numbers include at least the rational numbers, so that a unit function like *days* can assign fractions of days to events with very short but not instantaneous runtimes. That is, the runtimes of these events match $\lambda t[\text{days}(t) \leq 1]$ even though their runtime is not instantaneous. See Fox & Hackl (2006) for related discussion.

Lønning (1987) replaces the domain of numbers by more abstract entities, and considers which of the properties of the number scale must be imposed on these entities if one wants to avoid including all of mathematics in natural language. For example, if one wants to model the fact that people cannot always decide which of two very large numbers is bigger than the other, some numbers might be assumed not to be ordered with respect to one another. For ease of exposition, I do not follow Lønning's suggestion and I continue instead to use actual numbers in the mathematical sense, along with the total order \leq in which they stand.

2.5 Functions

As already shown in Figure 2.2, I assume that the domains of objects, events, intervals, degrees, and numbers presented in Section 2.4 are interconnected by a variety of relationships, which can be modeled as partial functions. Since all the functions in this book are partial, I use the word *functions* instead of the term *partial functions*. As described in this section, I assume that thematic roles map events to individuals; trace

functions map events to their locations in space and time; measure functions map individuals and events to degrees; and unit functions map degrees to numbers. In this section, I describe these functions in more detail. I have chosen those background assumptions that make these functions as similar as possible. In particular, I build on the concept of homomorphism, a special kind of function that preserves certain algebraic relationships such as sums and parts across domains. Homomorphisms constrain the structural relationships between domains and therefore, by extension, the domains themselves.

2.5.1 *Thematic roles (type $\langle v, e \rangle$)*

Thematic roles are semantic relations that represent different ways in which entities participate in events (Parsons 1990, Dowty 1991). There are two common views on thematic roles. On the traditional view, due to Gruber (1965) and Jackendoff (1972), thematic roles encapsulate generalizations over shared entailments of argument positions in different predicates. Thematic roles in this sense include *agent* (initiates the event, or is responsible for the event), *theme* (undergoes the event), *instrument* (used to perform an event), and sometimes also *location* and *time*. For example, if Brutus stabbed Caesar with a dagger in the Senate on March 15, then there is a stabbing event whose agent is Brutus, whose theme is Caesar, whose instrument is a dagger, whose location is the Senate, and whose time is March 15. I discuss location and time in Section 2.5.2. An alternative view sees thematic roles as verb-specific relations: Brutus is not the agent of the stabbing event but the stabber, Caesar is not its theme but its stabbee, and so on (Marantz 1984). This comes at the obvious price of missed generalizations, and it raises some technical problems. For example, it is not clear which role to assign to the subjects of coordinated sentences like *A girl sang and danced*. I adopt the more traditional, and relatively widespread, view that there is a small inventory of thematic roles that generalize over verbs. Since the semantic content of thematic roles is not an issue in this book, and since there does not seem to be a consensus on what to call the thematic role of the arguments of verbs like *own* or *be numerous*, I call the thematic role of the subject of any verb the agent, even if this is not always intuitive. Kratzer (1996) uses the word “holder” for the subject role of *own*, but she notes that this is only for convenience.

Unfortunately, there is no consensus on the inventory of thematic roles. It is difficult to identify these roles in borderline cases, though not impossible: a wide-coverage database of English verbs with their thematic roles, extending prior work by Levin (1993), was made available in Kipper-Schuler (2005).

I assume that thematic roles as semantic relations have syntactic counterparts which relate verbs to their arguments (see Section 2.10 for an illustration). This is a standard assumption in generative grammar, at least as far as the agent role is concerned: the “little *v*” head is assumed to relate verbs to their external arguments, which are usually their agents (Chomsky 1995). Other thematic roles can be taken to have syntactic

reflexes such as case marking and prepositions. Following Carnie (2006), I reserve the term *thematic role* for the semantic relation and I use the term *theta role* for its syntactic counterparts. However, some authors use these terms interchangeably.

The correspondence between theta roles and thematic roles is controversial. For example, Chomsky (1981) assumes that each argument is limited to one role. In many cases, though, it seems appropriate to assign more than one thematic role to the same argument: for example, the subject of a verb like *fall* can be regarded both as the agent and the theme of the event (Parsons 1990).

I adopt the assumption that each event has at most one agent, at most one theme, and so on. This assumption is generally called the Unique Role Requirement (URR) or thematic uniqueness. Arguments for the existence of thematic roles and for thematic uniqueness are reviewed in Carlson (1984, 1998), Parsons (1990), and Landman (2000). Formally, I assume that thematic roles are functions of type $\langle v, e \rangle$ rather than just relations of type $\langle v, et \rangle$. This choice is not essential because either option leads to type mismatches. I assume they are resolved by type-shifting (see Section 2.10).

Thematic uniqueness is a widespread but not universally accepted assumption. Krifka (1992) rejects it because, as he claims, one “can see a zebra and, with the same event of seeing, see the mane of the zebra as well” and one “can touch a shoulder and a person with the same event of touching.” However, nothing forces us to assume that there is just one rather than two events in each of these cases. Other arguments against thematic uniqueness are discussed and refuted in Landman (2000).

Thematic uniqueness can be integrated in a mereological setting where events are closed under sum formation. I assume for this purpose that thematic roles are their own algebraic closures (Krifka 1989b, 1998, Landman 2000). This property is also called *cumulativity* or *summativity of thematic roles*.

(34) Cumulativity assumption for thematic roles

For any thematic role θ and any subset E of its domain:

$$\theta(\bigoplus E) = \bigoplus (\lambda x \exists e \in E. \theta(e) = x)$$

This says that for any subset of the events on which a given thematic role θ is defined, we can compute the θ of their sum by summing up their θ s. A consequence of this assumption is that thematic roles are homomorphisms, or structure-preserving maps, with respect to the \oplus operation:

(35) Fact: Thematic roles are sum homomorphisms

For any thematic role θ :

$$\theta(e \oplus e') = \theta(e) \oplus \theta(e')$$

(The θ of the sum of two events is the sum of their θ s.)

What this says is that, for example, if e is a talking event whose agent is John and e' is a talking event whose agent is Mary, $e \oplus e'$ is an event whose agent is the sum of John and Mary. Thus, $e \oplus e'$ has a unique entity as its agent, even though this

entity is a proper sum. This sum operation is independent of time, so it may be that e took place two days ago and e' is taking place today. Discontinuous events of this kind will play an important role in the account of covert temporal distributivity in Section 8.6 and in the account of occasion readings of overt distributive items in Chapter 9.

As a reminder of the cumulativity assumption, I include the algebraic closure operator in the typographical representation of thematic roles, except in those cases where it is clear that they map events to atoms. For example, I generally write $\ast\text{agent}$ instead of agent , but I write $\text{Atom}(\text{agent})$ rather than $\text{Atom}(\ast\text{agent})$. This notation is meant to be reminiscent of a generalization of algebraic closure to partial functions (see Section 2.3.4).

It is sometimes useful to talk about the entity to which a certain event e is mapped under a thematic role θ , while abstracting over the identity of θ . I refer to this entity as the θ of e . For example, if x is the agent of e then x is the θ of e for $\theta = \text{agent}$.

Not everyone accepts the cumulativity assumptions for all thematic roles. Suppose with Kratzer (2003) that there are three events e_1, e_2, e_3 in which Al dug a hole, Bill inserted a rosebush in it, and Carl covered the rosebush with soil. Then, in virtue of these three events, one can say that there is also an event in which Al, Bill, and Carl planted a rosebush. Let e_4 be this event. Do we consider e_4 equal to the proper sum event $e_1 \oplus e_2 \oplus e_3$? If we do, this scenario is a counterexample to the cumulativity assumption, as Kratzer notes. The themes of e_1, e_2, e_3 are the hole, the rosebush, and the soil, and the theme of e_4 is just the rosebush. The theme of e_4 is not the sum of the themes of e_1, e_2 , and e_3 . This violates cumulativity.

I respond to this challenge by following my general strategy of not assuming that the mereological parthood relation should model all parthood relations that can be intuitively posited (see Section 2.3.1). In this case, I assume that e_4 is not actually the sum of e_1, e_2 , and e_3 . Even though the existence of e_4 can be traced back to the occurrence of e_1, e_2 , and e_3 , nothing forces us to assume that these events are actually parts of e_4 , just as we do not consider a plume of smoke to have among its parts the fire from which it comes, even though its existence can be traced back to the fire. Without the assumption that e_4 contains e_1 through e_3 as parts, Kratzer's objection against cumulativity vanishes. For more discussion, see Williams (2009).

2.5.2 Trace functions: runtime and location (type $\langle v, i \rangle$)

Trace functions map events to those intervals that represent their temporal and spatial locations. Since some events may not be situated in space and/or time, trace functions are partial. I assume the existence of two such functions, temporal trace or *runtime* and spatial trace or *location*. Following common usage in the literature, I also write these functions as τ and σ respectively. Since intervals correspond to spatiotemporal locations, trace functions do not simply indicate the amount of space and time that an event takes, but also the precise location in space and time. For example, suppose John

sings from 1pm to 2pm and Mary sings from 2pm to 3pm. Although each event takes the same time, their runtimes are different since they take place at different times. The temporal trace function mirrors this by mapping John's running and Mary's running to distinct intervals. These intervals are not directly related to each other, though they are mapped to the same numbers by the unit functions *hour*, *minute* and so on (see Section 2.5.4).

Following Link (1998a) and Krifka (1998), I assume that trace functions are sum homomorphisms: the runtime of the sum of two events is the sum of the runtimes. For example, in the scenario just mentioned, the runtime of the sum event that combines John's and Mary's running is the interval from 1pm to 3pm, and this is the sum of the intervals from 1pm to 2pm and from 2pm to 3pm. This is formally expressed in the equations in (36).

(36) **Trace functions are sum homomorphisms**

σ is a sum homomorphism: $\sigma(e \oplus e') = \sigma(e) \oplus \sigma(e')$

τ is a sum homomorphism: $\tau(e \oplus e') = \tau(e) \oplus \tau(e')$

(The location of the sum of two events is the sum of their locations, and similarly for their runtimes.)

This assumption is parallel to the cumulativity assumption on thematic roles described in Section 2.5.1. In general, trace functions are not very different from thematic roles. The main differences are that trace functions map events to intervals rather than to individuals, and that they are typically expressed by adjuncts rather than by arguments.

Trace functions cross the bridge between interval semantics and event semantics, since they allow us to realize the intuition that a predicate holds at a certain interval. Interval semantics assumes the existence of a relation *AT* that evaluates a proposition at an interval. It is useful to be able to refer to this relation also in event semantics, where instead of propositions we can relate event predicates to intervals. The following definition fixes the relation between *AT* and τ .

(37) **Definition: Holding at an interval**

$AT(V, i) \stackrel{\text{def}}{=} \exists e[V(e) \wedge \tau(e) = i]$

(An event predicate *V* holds at an interval *i* iff it holds of some event whose temporal trace is *i*.)

Since events are closed under sum, the homomorphism assumption has the consequence that two events whose runtimes are not adjacent have as their sum an event whose runtime is not continuous. See Section 2.5.4 for more discussion on noncontinuous intervals.

For more on temporal and spatial trace functions, see Hinrichs (1985), Lasnik (1988), Krifka (1998), Link (1998a), as well as Zwarts (2005, 2006), where the spatial and temporal trace functions are combined into one.

2.5.3 Measure functions (types $\langle e, d \rangle$ and $\langle v, d \rangle$)

I use the term “measure function” to denote a mapping between a class of physical objects and a degree scale that preserves a certain empirically given ordering relation, such as “be lighter than” or “be cooler than.” Typical measure functions are height, weight, speed, and temperature. It seems plausible to ascribe at least some of them to events rather than individuals. I have already mentioned the example due to Davidson (1969) of a sphere which can at the same time rotate quickly and heat up slowly. Provided that the rotating and the heating up are two separate events, and that *quickly* and *slowly* are event modifiers, we can avoid the undesirable conclusion that the sphere is both quick and slow by treating speed as a function that applies to events rather than individuals. See also Kratzer (2001) for an analysis that ascribes weights to states.

Some authors do not distinguish between trace functions and measure functions (e.g. Kratzer 2001). I prefer to treat them differently for two reasons. First, trace functions map into intervals and measure functions into degrees. Section 2.4.5 explains my reasons not to conflate these two categories. Second, the sum homomorphism assumption in (36) must be limited to trace functions (and thematic roles). Extending it to measure functions would have counterintuitive consequences, and would not be compatible with the standard conception of degrees as totally ordered entities. For example, suppose that we have two coins, c_1 and c_2 , which each weigh one gram. Following Lønning (1987), we represent this fact as follows: the measure function *weight* maps c_1 to its weight w_1 and c_2 to its weight w_2 , and the unit function *grams* maps both w_1 and w_2 to 1. If we wanted to extend assumption (36) to weight, it would follow that the weight function maps $c_1 \oplus c_2$, the sum of the two coins, to $w_1 \oplus w_2$. To model the fact that the two coins together weigh two grams, we assume that the function *grams* maps this entity to the number 2. From these assumptions, it follows that $w_1 \neq w_1 \oplus w_2$, that is, the weight of c_1 is not equal to the weight of $c_1 \oplus c_2$, which is plausible. From this fact it follows that $w_1 \neq w_2$, due to idempotence (Table 2.2). In other words, the two coins are mapped to different weight degrees, even though intuitively they “weigh the same.” More generally, whenever a measure function maps the sum of two objects to a different degree than it maps one of them, it also maps both objects to different degrees. As far as I can tell, no empirical reasons speak against this possibility.⁵ It is even consistent with the rest of the assumptions in this book, as long as one is willing to relax the standard assumption that degrees of the same sort are totally ordered (otherwise one needs to insist that w_1 precedes w_2 or vice versa). It seems conceptually unattractive to assume this proliferation of degree entities (but see Moltmann 2009 for an argument to the contrary, namely that totally ordered degrees should be replaced by entity-specific tropes). I find it easier to limit the sum homomorphism assumption to trace functions and to

⁵ I thank Roger Schwarzschild for discussing this issue with me.

exclude measure functions from it. Concretely, I assume that trace functions are sum homomorphisms but measure functions need not be. This motivates the conceptual distinction between trace functions and measure functions. However, this assumption is not essential.

2.5.4 Unit functions (types $\langle d, n \rangle$ and $\langle i, n \rangle$)

As discussed in Section 2.4.5, I adopt the proposal of Lønning (1987), according to which degrees occupy an intermediate layer between individuals and numbers. I refer to the functions that relate entities to degrees as *measure functions* and to the functions that relate degrees to numbers as *unit functions*. This use of the term *measure function* differs from Krifka (1998), whose framework lacks the intermediate layer of degrees. Krifka's measure functions map entities directly onto numbers (see Section 7.3). I use the term *degree* exclusively to refer to the things in the range of the measure functions, and I use the term *numbers* exclusively to refer to the things in the range of the unit functions.

As an illustration, suppose John weighs 150 pounds (68 kilograms) and measures 6 feet (183cm). These facts are represented as follows. The measure function *weight* maps John to a degree which the unit functions *pounds* and *kilograms* map to the numbers 150 and 68, respectively. The measure function *height* maps John to a degree which is mapped by the unit function *feet* to the number 6 and by the unit function *centimeters* to the number 183:

- (38) a. $\text{pounds}(\text{weight}(\text{john})) = 150$
 b. $\text{kilograms}(\text{weight}(\text{john})) = 68$
 c. $\text{feet}(\text{height}(\text{john})) = 6$
 d. $\text{centimeters}(\text{height}(\text{john})) = 183$

I assume that the total order of degrees mirrors the order of the range of the corresponding unit functions. For example:

- (39) $\text{height}(\text{john}) \leq \text{height}(\text{mary})$
 $\Leftrightarrow \text{meters}(\text{height}(\text{john})) \leq \text{meters}(\text{height}(\text{mary}))$

Splitting up the semantics of measurement into measure functions and unit functions introduces an additional degree of freedom in the semantic representation. This degree of freedom is needed to model the fact that pseudopartitives do not always specify the way in which the measure noun is related to the substance noun. For example, as Schwarzschild (2002) observes, a pseudopartitive like (40) can be understood as involving reference to an amount of oil that is either three inches deep, or (given the right context, e.g. when talking about a growing puddle), three inches in diameter. The two interpretations of the pseudopartitive can lead to different truth conditions. This can be modeled by the two interpretations in (40).

(40) three inches of oil

- a. $\lambda x[\text{oil}(x) \wedge \text{inches}(\text{height}(x)) = 3]$
- b. $\lambda x[\text{oil}(x) \wedge \text{inches}(\text{diameter}(x)) = 3]$

As we see, this difference can be expressed straightforwardly by Lønning's split: the interpretations differ in their measure function, but not in their unit function. Of course, not every pseudopartitive is ambiguous in this way. For example, (41) unambiguously involves reference to a quantity of oil whose volume is three liters (41a). It does not give rise to readings in which, say, the height of the oil is asserted to be three liters (41b). Since this reading is nonsensical, I assume that the unit function *liters* is a partial function that is only defined on degrees in the range of the measure function *volume* but not on those in the range of the measure function *height*. Many pseudopartitives are unambiguous as a consequence, even though in a pseudopartitive, only the unit function and not the measure function is overt.

(41) three liters of oil

- a. $\lambda x[\text{oil}(x) \wedge \text{liters}(\text{volume}(x)) = 3]$ OK
- b. $\lambda x[\text{oil}(x) \wedge \text{liters}(\text{height}(x)) = 3]$ *nonsensical*

For more motivation of the split between unit functions and measure functions, see Lønning (1987) and Schwarzschild (2002, 2006).

Following Krifka (1998), I assume that spatiotemporal intervals are related to numbers by functions like *hours* and *meters*. I call these functions unit functions as well, even though I make a formal difference between intervals and degrees. Intervals can be discontinuous (see Section 2.4.4). Degrees, by contrast, cannot (see Section 2.4.5).

The discontinuity of intervals poses a problem for the definition of unit functions. For example, if *a* is the interval from 2pm to 3pm, and *b* is the interval from 4pm to 5pm, then we have $\text{hours}(a) = \text{hours}(b) = 1$, but what about $\text{hours}(a \oplus b)$? Is it 3 or just 2? Let us call the first option the *generous construal* and the second option the *stingy construal* (see Kratzer 2007: n. 24). Evidence for either option is not hard to come by if we assume (see Section 3.3) that temporal *for*-adverbials use unit functions to map the runtimes of events to numbers. The relevant observations go back to Dowty (1979). For example, sentence (42) is coherent and favors the generous construal (see also Section 3.3 on *for*-adverbials with discontinuous intervals):

(42) John worked in New York for four years but he usually spent his weekends at the beach. (Dowty 1979: 334)

However, (43) can be true even if John served four non-consecutive one-year terms, and therefore favors the stingy construal.

(43) John served on that committee for four years. (Dowty 1979: 334)

The generous/stingy opposition should probably really be seen as a continuum, since the interpretation of following sentence appears to be intermediate between both construals:

- (44) From 2002 to 2008, John lived in New York for a total of four years, but he usually spent his weekends in Maine.

In the examples I consider in this book, stingy construals of *for*-adverbials do not often play a role. For this reason, I assume that unit functions are always interpreted generously. That is, I assume that unit functions map any discontinuous interval i to the same number as they map the smallest continuous interval containing i . In the example just mentioned, this means that $\text{hours}(a \oplus b) = 3$. Since $\text{hours}(a) + \text{hours}(b) = 2$, this assumption also means that unit functions are not sum homomorphisms.

Note that the present setup does not impose a choice between being generous and being stingy. We could equally well generalize unit functions to “unit relations,” and constrain them in such a way that both $\text{hours}(a \oplus b, 3)$ and $\text{hours}(a \oplus b, 2)$ hold in the above example. As discussed in Section 2.6.3, I assume that measure nouns have relational types like $\langle n, it \rangle$ and that their translations are based on unit functions. The meaning of measure nouns like *hour* could therefore easily be changed to accommodate unit relations:

- (45) a. $\llbracket \text{hour} \rrbracket$ (using unit functions) = $\lambda n \lambda t [\text{hours}(t) = n]$
 b. $\llbracket \text{hour} \rrbracket$ (using unit relations) = $\lambda n \lambda t [\text{hours}(t, n)]$

I do not pursue this route, however.

2.5.5 *The cardinality function (type $\langle e, n \rangle$)*

I assume that there is a partial function that maps sums which consist of singular individuals onto the number of singular individuals of which they consist. For example, the function maps a sum of three boys to the number 3. I assume that it is only defined on entities in the extension of count nouns. Following standard practice, I call this function *cardinality*, and I define it in terms of the set-theoretic notion of cardinality.

(46) **Definition: Cardinality**

For any sum x such that $*\text{Atom}(x)$, the *cardinality* of x , written $|x|$, is defined as $|\{y \mid y \leq_{\text{Atom}} x\}|$.

I assume that singular individuals are atomic (see Section 2.6.1). This means that singular individuals always have cardinality one. An alternative is found in Krifka (1989a), who assumes that there is a partial *natural unit* function, NU, which assigns plural individuals a number that intuitively represents their cardinality. NU is a primitive concept, so it does not require the atomicity assumption.

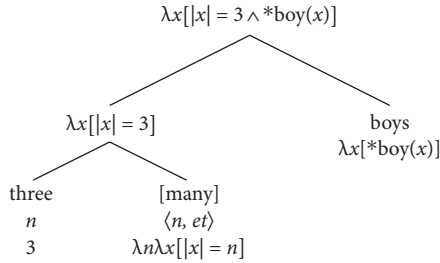


FIG. 2.3 LF of the noun phrase *three boys*.

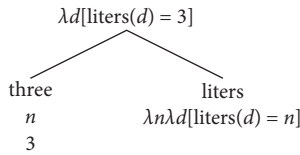


FIG. 2.4 LF of the noun phrase *three liters*.

In order to be able to give a single lexical entry to numerals whether they occur in measure phrases or in run-of-the-mill noun phrases, I give number words the type n and I let them denote actual numbers, as in Krifka (1989a), Hackl (2001), and Landman (2004). For example, *three* denotes the number 3 rather than a generalized quantifier or a predicate over pluralities with three atoms. To represent numeral noun phrases, I assume that a silent head [many] introduces the cardinality function, as in Figure 2.3 (see Hackl 2009 for a similar proposal). This represents the sum reading of the numeral *three boys* (see Section 2.8). The lexical entry for *boys* is justified in Section 2.6.2. In measure phrases, the number word combines with the measure noun, which is assumed to be of type $\langle n, dt \rangle$ (see Section 2.6.3), as shown in Figure 2.4. In this case, the [many] head is not necessary.

In my entries for [many] and for unit functions, I use an *exactly* interpretation rather than an *at least* interpretation. That is, I use an equality sign rather than a \geq sign. This decision keeps things compatible with Krifka (1998), who assumes that predicates like *eat three apples* are quantized and therefore do not apply to events in which more than three apples are eaten. There is no consensus in the literature on whether the literal meaning of *three boys* is *at least three boys* (e.g. Horn 1972, 1989, Barwise & Cooper 1981, Levinson 2000, van Rooij & Schulz 2006) or *exactly three boys* (e.g. Harnish 1976, Sadock 1984, Partee 1987, Carston 1998, Geurts 2006, Breheny 2008). For a review of this literature, see Nouwen (2006) and Kennedy (2009). In an event semantics, it does not matter much whether numerals are given an *exactly* or an *at least* reading, since the maximality conditions involved in *exactly* readings have to be realized through other mechanisms anyway (Krifka 1999, Landman 2000, Robaldo 2011, Brasoveanu 2013). (In this book, I only account for sentences involving unmodified numerals like

three boys. Modified numerals like *exactly three boys* and *at most three boys* can also be handled in principle, but they raise additional issues involving event maximality (Krifka 1989a, 1999, Brasoveanu 2013).)

2.6 Nouns

In this section, I describe my assumptions regarding the count/mass opposition, the singular/plural opposition, and the status of measure nouns and group nouns.

I analyze nouns as one-place predicates over entities, or equivalently, as denoting sets of entities, except for measure nouns, which I assume to denote unit functions. The entities in the denotations of most nouns are ordinary objects (Section 2.4.1), but I assume that derived nouns like *running* and, by extension, pseudopartitives like *three hours of running* involve reference to events.

Following standard usage in the literature, I assume that the terms *mass noun* and *count noun* are defined with reference to morphological and syntactic properties rather than semantic properties. In English, mass nouns are compatible with the quantifiers *much* and *little* and reject quantifiers such as *each*, *every*, *several*, *a/an*, *some* and numerals (Bunt 2006, Chierchia 2010). Most mass nouns are incompatible with plural morphemes, and in the absence of these morphemes only mass nouns are able to form noun phrases by themselves.

It is well known that many nouns can be used as count nouns or as mass nouns:

- (47) a. Kim put an apple into the salad.
b. Kim put apple into the salad.

This raises the issue of whether the two occurrences of *apple* represent two different words or two different senses of one word. The differences between these positions are discussed in Pelletier & Schubert (2002). I assume that the two occurrences correspond to two different lexical entries, one singular count noun and one mass noun. The denotations of these entries are taken to have different formal properties, as described in the following sections. For example, the former entry only applies to mereological atoms, but the latter need not.

Although my discussion is phrased in terms of nouns, similar considerations also apply to nominals (e.g. *drinkable water* and *red barn* are nominals). The assumptions I make about the meanings of nouns carry over to nominals. This means that instead of talking about count and mass nouns, more exact terms would be count and mass expressions (Pelletier & Schubert 2002). However, for simplicity I use the term *nouns* rather than “nominals” or “expressions.”

2.6.1 Singular count nouns

I adopt the standard assumption that singular count nouns denote sets of individuals. For example, *table* denotes the set of all tables. I call the entities in the denotation

of a count noun *singular individuals*. As described in Section 2.3.1, I assume that all singular individuals are pure mereological atoms. The notion of mereological part must be distinguished from the intuitive notion of part. For example, the leg of a table is not a mereological part of the table. Since all entities in the denotation of a count noun are atoms, all count nouns have quantized reference (Section 2.3.5).

This assumption is standard in accounts that are not primarily concerned with the count/mass distinction, but it is not self-evident (Kratzer 2016, Casati & Varzi 1999: 112–15). It has been challenged by Zucchi & White (2001) on the basis of words like *twig*, *rock*, and *sequence*. Intuitively, a twig may have a part that is again a twig, a rock may have a part that is again a rock, and so on. Following my general stance described in Section 2.3.1, I assume that the relation between a twig and its part is not mereological parthood, just like we are not committed to the assumption that the relation between John and his arm is mereological parthood. Regarding *twig*-type nouns, I assume that the meaning of these nouns is partially specified by context, and that when the context is fixed, each of these nouns denotes a quantized set. This assumption seems justified given that these nouns can be used with numerals and quantifiers just like other count nouns (*a twig*, *two rocks*, *every sequence*). An example of a formal implementation of this assumption can be found in Chierchia (2010: sect. 5.2), with respect to the meaning of the similar noun *quantity*. For more discussion, see Rothstein (2010).

2.6.2 Plural count nouns

I assume that plural count nouns denote sets of entities, and that a plural count noun is the algebraic closure of its singular counterpart. These assumptions are common but not uncontroversial. Different analyses of plural count nouns result in different answers to the questions of whether their denotations include singular entities and whether they include proper parts of singular entities. As we will see, dependent plurals pose a special challenge for any analysis, because they show that plural count nouns are interpreted differently depending on the polarity of their context.

In most mereological approaches to the semantics of count nouns, a singular noun denotes a quantized set, and a plural count noun denotes either the algebraic closure of the set denoted by its singular form, or a certain subset of this closure. This is a controversial point. The issue is whether the denotation of a plural noun also contains entities denoted by its singular form. In the terminology of Farkas & de Swart (2010), the two views on this question are termed “exclusive” and “inclusive.” Their own view is a combination of the two and can be called mixed. These three views can be summarized as follows, and are illustrated in Figure 2.5.

- *Exclusive view*: Singular and plural forms of a count noun *N* denote disjoint sets. The plural form denotes the algebraic closure of its singular form with the singular individuals removed, as described by the equation in (48). The plural

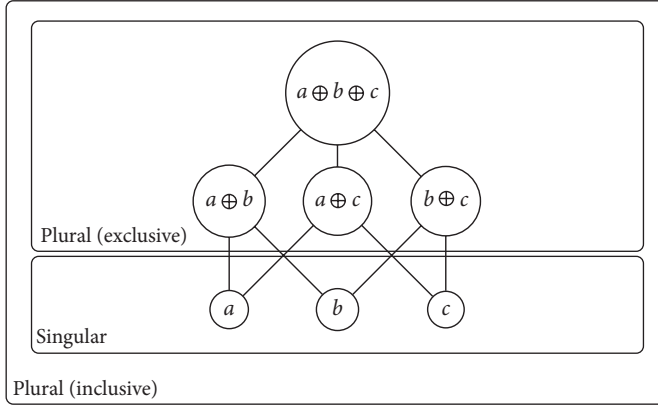


FIG. 2.5 Different views on the plural.

form N_{pl} essentially means the same as *two or more N*. For example, the plural count noun *apples* applies to sums consisting of two or more apples (Link 1983; Chierchia 1998a).

$$(48) \llbracket N_{pl} \rrbracket = * \llbracket N_{sg} \rrbracket - \llbracket N_{sg} \rrbracket$$

$$(49) \text{ a. } \llbracket \text{cat} \rrbracket = \{a, b, c\}$$

$$\text{ b. } \llbracket \text{cats} \rrbracket = * \llbracket \text{cat} \rrbracket - \llbracket \text{cat} \rrbracket = \{a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c\}$$

- *Inclusive view*: This is the view I adopt. The plural form of a count noun N denotes a proper superset of the singular form, namely its algebraic closure, as described by the equation in (50). The plural form essentially means the same as *one or more N* (Krifka 1989b; Sauerland 2003; Sauerland, Anderssen & Yatsushiro 2005; Spector 2007; Chierchia 2010). For example, the plural count noun *apples* applies to sums consisting of one or more apples. When singular reference is intended, singular and plural forms are in competition, and inclusive theories usually appeal to pragmatic notions in order to explain why the singular form blocks the plural form.

$$(50) \llbracket N_{pl} \rrbracket = * \llbracket N_{sg} \rrbracket$$

$$(51) \text{ a. } \llbracket \text{cat} \rrbracket = \{c_1, c_2, c_3\}$$

$$\text{ b. } \llbracket \text{cats} \rrbracket = * \llbracket \text{cat} \rrbracket = \{c_1, c_2, c_3, c_1 \oplus c_2, c_2 \oplus c_3, c_1 \oplus c_3, c_1 \oplus c_2 \oplus c_3\}$$

- *Mixed view*: The plural form of a count noun N is ambiguous between the meanings it has on the exclusive and on the inclusive view (Farkas & de Swart 2010).

If singular count nouns are taken to apply to mereological atoms, the exclusive view holds that plural count nouns apply to proper sums, and the inclusive view holds that they apply both to atoms and to proper sums.

Environments such as downward-entailing contexts and questions present a challenge for the exclusive view because in these environments, plurals can be verified by singular individuals (for this argument and its sources, see Schwarzschild 1996: 5). For example, suppose that there is only one doctor. On the exclusive view, the plural form *doctors* denotes the empty set, from which it is not possible to compositionally retrieve any information about the intended referent (the doctor). Therefore, the exclusive view has major problems deriving the meanings of sentences like (52) compositionally.

- (52) a. No doctors are in the room.
b. Are there doctors in the room?

Bare plurals can occur in a particular configuration in which their semantic contribution cannot be modeled completely either by the inclusive view or by the exclusive view. In this configuration, which has been studied extensively by Zweig (2008, 2009), the bare plural occurs as a verbal argument and is c-commanded by another argument or adjunct of the same verb whose head noun is typically plural as well. Consider for example the bare plural *kites* in the examples in (53), adapted from Zweig (2008).

- (53) a. Boys flew kites.
b. Some boys flew kites.
c. Several boys flew kites.
d. Five boys flew kites.

These sentences have a reading in which the semantic contribution of the bare plural *kites* consists of two components. First, each of the boys in question flew one or more kites. I call this the *distributivity component*. Second, the total number of kites flown was two or more (though as Cohen (2005) argues, the total number implicated may be even higher than two). Loosely following Zweig, I call this the *multiplicity component*. Bare plurals which contribute these two components are called *dependent plurals* (de Mey 1981).

An inclusive analysis by itself would wrongly predict that *kites* should be synonymous with *one or more kites*. In (53), this prediction does not explain the multiplicity component. An exclusive analysis would predict that *kites* is synonymous with *two or more kites*, and it would therefore make the right predictions in sentences with dependent plurals like (53).⁶ However, we have already seen above that an exclusive analysis is not viable for bare plurals in downward-entailing contexts. As Zweig shows, these bare plurals never contribute a multiplicity component, even when they occur in the same syntactic position as dependent plurals. For example, the bare plural *kites* in the sentences in (54) is synonymous with *one or more kites*, because their subjects quantify over boys who flew one or more kites. This is contrary to the

⁶ Zweig assumes that the subjects of all sentences with dependent plurals are interpreted in situ and do not distribute over their verb phrases. That bare plurals are interpreted distributively in the sentences considered here is a result of verb semantics and meaning postulates. See Section 4.5.1.

exclusive analysis, according to which they should quantify over boys who flew two or more kites.

- (54) a. No boys flew kites.
b. Few boys flew kites.

In a nutshell, the puzzle is that sentences like those in (53) are incompatible with an inclusive analysis, while sentences like those in (52) and (54) are incompatible with an exclusive analysis. However, bare plurals are not simply ambiguous between an inclusive and an exclusive reading, because all these sentences are unambiguous.

I adopt the solution to this puzzle proposed in Spector (2007) and Zweig (2008, 2009): the exclusive inference is a grammaticalized scalar implicature. Scalar implicatures only surface when they strengthen the meaning of the sentence. The truth-conditional meaning of a bare plural is inclusive and provides the distributivity contribution, and its exclusive part is a scalar implicature. Following Chierchia (2006), scalar implicatures can be added to truth-conditional meanings, but only if this strengthens the truth conditions of the entire sentence. This is the case in upward-entailing contexts: for example, *five boys flew two or more kites* is stronger than *five boys flew one or more kites*, so the former meaning is retained. But it is not the case in downward-entailing contexts: *no boys flew two or more kites* is weaker than *no boys flew one or more kites*, so the latter meaning is retained. In this way, bare plurals are interpreted exclusively in upward-entailing contexts and inclusively in downward-entailing contexts.

2.6.3 Measure nouns

I assume that measure nouns like *liter*, *kilogram*, and *year* have a separate analysis from other count nouns, namely, they have the type $\langle n, dt \rangle$ (relations between degrees and numbers) or $\langle n, it \rangle$ (relations between intervals and numbers) and their denotation makes reference to unit functions like *liters* and *kilograms* in the sense of Section 2.5.4. I assume that singular and plural forms of measure nouns have identical denotations:

- (55) a. $\llbracket \text{liter} \rrbracket = \llbracket \text{liters} \rrbracket = \lambda n \lambda d [\text{liters}(d) = n]$
b. $\llbracket \text{year} \rrbracket = \llbracket \text{years} \rrbracket = \lambda n \lambda t [\text{years}(t) = n]$

I conjecture that my analysis of measure nouns could also be extended to collection nouns like *bunch*, and that the ontological difference between individuals and unit functions might provide a handle on the problem of distinguishing collection nouns from group nouns. See Pearson (2009) and Section 2.6.4.

In connection with my assumption that number words denote numbers (Section 2.4.6), this translation of measure nouns has the consequence that measure phrases like *three liters* have predicative type, similarly to other weak quantificational noun phrases such as *three boys* (Section 2.4.6). Different versions of the predicative analysis of measure phrases are defended in Zwarts (1997) and Schwarzschild (2006), among others.

Motivation for this analysis of measure nouns comes from the ambiguity between measure and individuating readings of container pseudopartitives (Rothstein 2009 and references therein):

- (56) three glasses of wine
 a. *Measure reading*: a quantity of wine that corresponds to three glassfuls
 b. *Individuating reading*: three actual glasses containing wine

Only the individuating reading, but not the measure reading, entails the existence of actual glasses. This fact can be explained if certain nouns like *glass* are taken to be ambiguous between an ordinary and a measure noun interpretation. The measure noun interpretation can be taken to give rise to the measure reading and the latter to the individuating reading:

- (57) three glasses of wine
 a. *Measure reading*: $\lambda x[\text{wine}(x) \wedge \text{glasses}(x) = 3]$
 (a quantity of wine that corresponds to three glassfuls)
 b. *Individuating reading*: $\lambda x[|x| = 3 \wedge * \text{glass}(x) \wedge \text{contains}(x, \text{wine})]$
 (three actual glasses containing wine)

In the individuating reading, *wine* appears as the argument of the relation *contains*. Rothstein treats *wine* as a kind in this case (see Section 2.4.2). Another possibility would be to treat it as the set of all wine entities or its characteristic function, corresponding to its occurrence in the measure reading. Since I ignore individuating readings, I leave this issue open.

In contrast to container pseudopartitives, measure pseudopartitives like *three liters of water* only have the measure reading, which indicates that the noun *liter* only has a measure noun interpretation. In this book, I concentrate on measure pseudopartitives, since I am not interested in individuating readings.

2.6.4 Group nouns

Group nouns are nouns like *committee*, *army*, and *league*. Barker (1992) defines group nouns as count nouns that can take an *of*-phrase containing a plural complement but not a singular complement (*the group of armchairs/*armchair*, *a committee of women/*woman*, *an army of children/*child*, etc.). By contrast, non-group nouns either take no *of*-complements (**a book of page/pages*) or take *of*-complements which can be singular and/or plural (*a piece of cookie/*cookies*, *a picture of {a horse / horses}*).

Although at first sight, we might not be ready to view the individuals in the denotation of group nouns as atomic individuals, I assume that the entities in the denotation of singular group nouns are indeed mereological atoms, like other singular count nouns. This analysis is defended in Barker (1992), Schwarzschild (1996), and Winter (2001) against previous proposals such as Bennett (1974), in which group nouns

involve reference to pluralities. I also follow Barker (1992) in assuming that there is no semantic operation that would allow us, for example, to recover the members of a committee from the atoms in the denotation of the noun *committee*. In particular, the relation between a committee and its members is not the same as the relation between the sum and group interpretations of a plural definite like *the boys* (see Section 2.8). In terms of Section 2.3.1, I assume that the entities in the denotation of a group noun are pure atoms.

I use the name *group partitives* for the *of*-constructions that Barker uses in his test. Group partitives are superficially similar to pseudopartitives like *a liter of water* or *a box of cookies*. According to Barker's test, we might classify container words like *box* as group nouns (*a box of cookies*/**cookie*). To the extent that we can always find (singular) mass nouns that can go together with these container words (*a box of jewelry*, *a box of paperwork*), this will not be a problem. I am not sure if this is always possible, though. My account of pseudopartitives is not intended to apply to group partitives. I propose to distinguish them from each other as follows. Pseudopartitives based on container words like *three glasses of wine* are ambiguous between a measure reading, in which only the quantity of the wine is reported, and an individuating reading, which involves reference to actual glasses (56). Following Rothstein (2010), I assume that the two readings correspond to different syntactic configurations (see Section 2.6.3 for details). Aside from metaphorical uses such as *an army of ants*, group partitives only have an individuating reading (58).

(58) a committee of women

- a. *Measure reading (unavailable)*: a set of women whose number corresponds to the membership of a committee
- b. *Individuating reading*: an actual committee whose members are women

To avoid the problem caused by pseudopartitives, I propose to extend Barker's diagnostic for group nouns as follows: a group noun *N* is a count noun that can take an *of*-phrase containing a plural complement but not a singular complement, provided that the resulting interpretation entails the presence of actual *N*s. This prevents *box* from being incorrectly diagnosed as a group noun, because *a box of cookies* has a measure reading that does not entail the presence of an actual box.

This amended diagnostic faces a potential problem with nouns like *bunch* and *pile*, for whose referents it is difficult to establish independent existence criteria. As Pearson (2009, 2010) shows, these nouns behave differently from what we might call "true" group nouns like *committee* and *army*. Pearson calls the former *collection nouns* and the latter *committee nouns*. I conjecture that the *of*-constructions in which collection nouns appear are pseudopartitives and not group partitives. I leave open whether the predictions in this book extend to pseudopartitives with collection nouns.

2.6.5 Mass nouns

Following Bealer (1979) and Krifka (1991), we can distinguish three analyses of mass nouns. According to the *general term analysis*, the denotation of a mass noun is a set of entities. For example, on the general term analysis, the mass noun *gold* denotes the set of all gold entities, similarly to a count noun. According to the *singular term analysis*, the denotation of a mass noun is a sum. On the singular term analysis, *gold* denotes the sum of all gold entities, similarly to a proper name. A variant of the singular term analysis is the *kind reference analysis*, on which the mass noun *gold* denotes the kind *gold* (see Section 2.4.2). Some authors also advocate a dual analysis according to which mass nouns are systematically ambiguous between a reading on which they denote sums or kinds, and a reading on which they denote sets.

I assume that all mass nouns can be translated as sets of entities. This is compatible with both the general term analysis and the dual analysis. I do not assume that mass terms have divisive reference, because of the problem of fake mass nouns like *furniture* (see Section 2.4.1). However, I assume that any entity in the denotation of a mass noun N that is larger than a certain threshold $\varepsilon(C)$ can be divided into parts which are again in the denotation of N . This idea is sketched in Krifka (1989b) and can be formalized as follows:

(59) **Definition: Approximate divisive reference**

$$\text{ADIV}_{\varepsilon(C)}(P) \stackrel{\text{def}}{=} \forall x[P(x) \wedge \neg \varepsilon(C)(x) \rightarrow \forall y[y < x \wedge \neg \varepsilon(C)(y) \rightarrow P(y)]]$$

(A predicate P is approximately divisive with respect to a threshold $\varepsilon(C)$ iff whenever it holds of something, it also holds of each of its proper parts, excluding entities below the threshold.)

I assume that the threshold is set depending on a context C , and that in the case of fake mass nouns it is large enough to exclude putatively atomic entities like pieces of furniture. Approximate divisive reference avoids the problem of minimal parts, which is treated in detail in Chapter 5.⁷

2.6.6 Kind-referring readings

So far, I have concentrated on the object-referring reading of nouns. Let me briefly comment on another reading, the so-called kind-referring reading (Krifka, Pelletier, Carlson, ter Meulen, Chierchia & Link 1995, Krifka 2004, Delfitto 2005). Both readings occur in bare NP uses of count and mass nouns and can be contrasted in (60) (Milsark 1974, Carlson 1989):

⁷ Anticipating the discussion in that chapter, if a mass noun has approximate divisive reference with respect to a threshold $\varepsilon(C)$, then for any dimension d , it has universal stratified reference with respect to $\varepsilon(C)$ and d .

- (60) Typhoons arise in this part of the Pacific.
- a. *Kind-referring reading*: It is a property of typhoons in general that they arise in this part of the Pacific.
 - b. *Object-referring reading*: This part of the Pacific has the property that some typhoons arise in it.

Existing analyses of the kind/object opposition can be ordered in three categories, similarly to the different analyses of mass nouns (Krifka 1991, Krifka, Pelletier, Carlson, ter Meulen, Chierchia & Link 1995, Lasersohn 2011):

- (a) all bare NPs denote predicates over entities;
- (b) all bare NPs denote special entities, such as maximal sums or kinds;
- (c) kind-referring NPs denote entities, object-referring NPs denote sets of entities, and the two are related by type-shifting operators.

The ambiguity in (60) suggests that two uses of NPs must be distinguished semantically, as expected on analysis (c). For the purpose of this book, in analogy to my assumptions about mass nouns, I assume that all plural count nouns can be translated as sets of entities. This is compatible with both analyses (b) and (c).

The question of how to model the kind/object opposition touches on my account of substance nouns in pseudopartitives and of dependent plurals. I assume that substance nouns in pseudopartitives are object-referring, since on this view they denote sets, which makes the formulation of strata theory easier. This is perhaps the prevailing view, though not the only one: a kind-referring analysis of pseudopartitives is found in Ionin, Matushansky & Ruys (2006). In the case of dependent plurals, I follow Zweig (2008, 2009) in adopting an object-referring analysis, as described above.

2.7 Verbs

In this section, I describe my assumptions regarding the semantics of verbs, which I analyze as one-place predicates over events. I assume that the denotations of all verbs, but not all verb phrases, are closed under sum formation.

2.7.1 *The Neo-Davidsonian position and its alternatives*

There is no agreement in the formal semantic literature on how to represent the meanings of verbs. Early work, like some modern authors, simply represents the meaning of a verb with n syntactic arguments as an n -ary relation. A transitive verb, for example, is assumed to denote a two-place relation. Against this, Davidson (1967) argued that verbs denote relations between events (see Section 2.4.3) and their arguments, so that a transitive verb denotes a three-place relation. Once events have been introduced, we can express the relationship between events and their arguments by thematic roles (see Section 2.5.1). This is the so-called Neo-Davidsonian position

TABLE 2.3. Approaches to verbal denotations

Position	Verbal denotation	Example
Traditional	$\lambda y \lambda x [\text{stab}(x, y)]$	$\text{stab}(b, c)$
Classical Davidsonian	$\lambda y \lambda x \lambda e [\text{stab}(e, x, y)]$	$\exists e [\text{stab}(e, b, c)]$
Neo-Davidsonian	$\lambda e [\text{stab}(e)]$	$\exists e [\text{stab}(e) \wedge \text{agent}(e, b) \wedge \text{theme}(e, c)]$
Asymmetric	$\lambda y \lambda e [\text{stab}(e, y)]$	$\exists e [\text{agent}(e, b) \wedge \text{stab}(e, c)]$

(e.g. Carlson 1984, Parsons 1990, Schein 1993). Finally, Kratzer (2000) argues for an asymmetric position, according to which only agents are represented as thematic roles. Table 2.3 illustrates how the different positions are reflected in analyses of the sentence *Brutus stabbed Caesar*.

There does not seem to be a best practice of how to implement event semantics and how to combine verbal projections with their arguments. Let me briefly mention some empirical and theoretical motivations that lead me to adopt Neo-Davidsonian semantics over these approaches. In Neo-Davidsonian event semantics, thematic roles can be treated as of one and the same kind as other event properties that encode their location in space and time. This makes it a natural fit for the parallel treatment of distribution over individuals and distribution over time intervals. Furthermore, Neo-Davidsonian semantics allows us to adopt a uniform semantic architecture that makes it easier to express the theories in this book. For example, it exposes thematic roles to the compositional semantics, as opposed to keeping some or all of them implicit within the lexical entry of the verb. As we will see in Chapter 8, this makes it easier to access them via indices or function application. It also allows us to give the same type to all verbal projections, which makes it easier for distributivity operators to apply uniformly to each projection. Moreover, event semantics makes it possible to treat verbal and nominal projections as predicates of the same arity. This is useful from the point of view of algebraic semantics because it allows us to formulate cross-domain generalizations more easily. Finally, it allows us to treat distance-distributive items uniformly no matter if they occur within verbal or within nominal projections (see Chapter 9). That said, it may be possible to reformulate some of the content of this book in other approaches than Neo-Davidsonian event semantics if one wishes to do so. For a detailed study of what it takes to reformulate Neo-Davidsonian theories in classical Davidsonian and eventless frameworks, see Bayer (1997). It may also be possible to work in a framework that exposes some but not all thematic roles to the compositional process (e.g. Kratzer 1996). This would lead one to expect asymmetries between the agent role and other thematic roles. I have argued elsewhere that the asymmetries that Kratzer tries to model correlate with syntactic positions and not with thematic roles (Champollion 2010a). For more discussion of Kratzer's asymmetries, see Williams (2009).

Since the identity of thematic roles is not important for my purposes, I make the simplifying assumption that the subject of a sentence in active voice is always its agent, as mentioned in Section 2.5.1. Although I assume that the predicates denoted by verbs range over events and not over individuals, I say that a verb “applies” to an individual as a shorthand for stating that the individual is the agent of an event to which the verb applies. This is simply a matter of convenience, and should not be confused with adopting a traditional or classical Davidsonian position.

2.7.2 *Lexical cumulativity*

In Section 2.6.1, I called individuals in the denotation of singular nouns *singular individuals*. Similarly, I call an event whose thematic roles map it exclusively to singular individuals a *singular event*, and an event that is not a singular event a *plural event*. Choosing these terms allows me to sidestep the question of whether singular events should be represented as mereological atoms. Different views exist on this question. Landman (2000) assumes that all singular events are atoms because this allows him to treat distributivity and plurality as the same thing. For Krifka (1998), this is not the case. For example, the event in which John reads a certain book has proper parts in which John reads part of the book. In fact, Krifka leaves open whether atomic events exist at all. (The discrepancy between the two accounts is likely related to the fact that Landman does not try to model the aspectual phenomena that motivate Krifka’s assumptions.) I take Krifka’s view and make no assumptions about whether singular events are atomic.

In algebraic event semantics, the sum of any two events is itself an event (see Section 2.4.3). In general, the sum of two singular events is a plural event. For example, let e_1 be the event in which John (j) lifts a certain box b and e_2 the event in which Mary (m) lifts a certain table t . The sum $e_1 \oplus e_2$ is itself an event. The agent of e_1 is j and the agent of e_2 is m . Given that thematic roles are sum homomorphisms (see Section 2.5.1), the agent of the sum event $e_1 \oplus e_2$ is $j \oplus m$, the sum of their agents. Since the entity $j \oplus m$ is not a singular individual, the event $e \oplus e_2$ is a plural event.

While it might seem intuitive to assume that verbs only apply to singular events, I take the opposite view: verbs may apply both to singular and to plural events. More specifically, whenever two events are in the denotation of a verb, so is their sum. This is a common and well-motivated assumption (Scha 1981, Schein 1986, 1993, Lasersohn 1989, Krifka 1989b, 1992, Landman 1996, 2000). Following Kratzer (2007), I call this assumption *lexical cumulativity*. Within a Neo-Davidsonian framework, lexical cumulativity is complemented by the analogous assumption that all thematic roles are closed under sum formation (see Section 2.5.1). In the context of other approaches to verbal denotations where verbs denote relations, lexical cumulativity can be implemented as assuming that these relations are closed under pointwise sum formation (see Section 2.3.1).

In the previous example, lexical cumulativity has the consequence that the verb *lift* applies not only to the event e_1 and to the event e_2 , but also to their sum $e_1 \oplus e_2$. In general, verbs can be said to have plural denotations, in the sense that their denotations obey the same equation (61) as plural count nouns on the inclusive view (50), repeated in (62):

$$(61) \quad \llbracket V \rrbracket = * \llbracket V \rrbracket$$

$$(62) \quad \llbracket N_{pl} \rrbracket = * \llbracket N_{sg} \rrbracket$$

Following Kratzer (2007), I include the algebraic closure operator in the typographical representation of verb meanings as a reminder of the lexical cumulativity assumption. For example, instead of writing $\lambda e[\text{lift}(e)]$ for the meaning of the verb *lift*, I write $\lambda e[*\text{lift}(e)]$.

The lexical cumulativity assumption is motivated by the entailments in (63) and (64) (Krifka 1989a, 1992). Because of the parallelism between (61) and (62), the explanation of these entailments is completely analogous to the explanation of the entailment in (16), repeated here as (65), which motivated the treatment of plurality in Link (1983).

- (63) a. John slept.
 b. Mary slept.
 c. \Rightarrow John and Mary slept.
- (64) a. John lifted box b .
 b. Mary lifted table t .
 c. \Rightarrow John and Mary lifted box b and table t .
- (65) a. John is a boy.
 b. Bill is a boy.
 c. \Rightarrow John and Bill are boys.

As Kratzer (2007) notes, another motivation for lexical cumulativity comes from iterative interpretations of verbs. In English, semelfactive verbs such as *cough* and achievements such as *find* are systematically ambiguous or underspecified between a “punctual” and an iterative sense. (I use the term “punctual” with caution because I do not want to suggest that it takes no time at all to cough or to find something. Instants are not part of my temporal ontology anyway, see Section 2.4.4.) For example, the sentence *John coughed* can be understood as saying that John coughed once, or that he coughed several times (Carlson 2006), and the sentence *John found fleas* can be understood as saying that there was one punctual event in which John found some fleas, or that he found them over an ongoing period of time, possibly one by one. The lexical plurality assumption means that a predicate like $\lambda e[*\text{cough}(e) \wedge * \text{agent}(e) = j]$ applies both to singular and to plural coughing events. This fact captures the systematic ambiguity of semelfactive predicates.

Iterative interpretations of semelfactives typically involve an alternation of times at which the predicate is true and times at which it is false. The lexical cumulativity assumption does not semantically entail this alternation, because sum events with continuous runtimes fall under the denotation of the event predicate as well as sum events with discontinuous runtimes. This is justified if the alternation is a pragmatic implicature, as opposed to a semantic entailment, as argued by Egg (1995). For more on iterativity, see Lasersohn (1995) and van Geenhoven (2004). These authors model iterativity by dedicated verb-level iterativity operators, but these operators are more difficult to motivate independently than lexical cumulativity (Kratzer 2007). Since iterativity appears to be possible with all English predicates, an operatorless approach is *ceteris paribus* preferable to an operator-based approach, as Kratzer notes.

While lexical cumulativity does not entail that all verb phrases have cumulative reference (e.g. the sum of two events in the denotation of the verb phrase *carry exactly two pianos* is not again in its denotation, because it involves four rather than two pianos), there are some verb phrases that do obtain cumulative reference as a consequence of lexical cumulativity. Among cumulative verb phrases we find all those with constant objects and those with cumulative objects. For example, lexical cumulativity has the effect that the lexical predicate *carry* has cumulative reference, and applies not only to singular carrying events but also to sums of such events. The effect of this assumption on the denotation of the verb phrase *carry the piano* is that it denotes a predicate which applies not only to singular carrying events whose theme is the piano, but also to sums of such events, as in (66a). As a consequence, the entire phrasal predicate *carry the piano* has cumulative reference, and it has the same denotation as it would have if algebraic closure were applied to the entire phrasal predicate, as in (66b). In other words, (66a) and (66b) are equivalent.

- (66) a. $\lambda e[*\text{carry}(e) \wedge *\text{theme}(e) = ix.\text{piano}(x)]$
 b. $*\lambda e[\text{carry}(e) \wedge *\text{theme}(e) = ix.\text{piano}(x)]$

The concept of lexical cumulativity is a theoretical assumption which is implemented by adopting the premise that every verb denotes a predicate that satisfies cumulative reference. This implementation of lexical cumulativity affects the denotations of all verbs without exception. A consequence of this assumption is that it is not possible to use cumulative reference to model the telic/atelic opposition, as is sometimes proposed (e.g. Egg 1995, Zwarts 2005). Lexical cumulativity entails that even certain telic predicates have cumulative reference, such as achievement verbs with definite objects. For example, *reach the summit* is telic but cumulative.

2.8 Noun phrases

I assume a treatment of nondistributive noun phrases along the lines of Landman (1996, 2000). Distributive noun phrases such as *every boy* and *each girl* are discussed in detail in Section 9.5.

In Landman's account, definite and indefinite noun phrases are ambiguous between sum and group interpretations. This ambiguity is motivated by cumulative and collective readings.

Cumulative readings involve two plural entities A and B and a relation R that holds between the members of these plural entities in a certain way. In canonical examples of cumulative readings, A and B are introduced by two plural definite or indefinite arguments of a verb that is distributive on both these arguments, and R is introduced by this verb. A cumulative reading with a distributive predicate licenses the inference that the relation R relates each singular individual in A to at least one singular individual in B, and vice versa. For example, (67a) involves reference to a plural individual consisting of 600 Dutch firms (A), another one consisting of 5,000 American computers (B), and a relation of using (R). The most prominent reading of (67a), paraphrased in (67b), is a cumulative reading.

- (67) a. 600 Dutch firms use 5,000 American computers. (Scha 1981)
 b. 600 Dutch firms each use at least one American computer and 5,000 American computers are each used by at least one Dutch firm.

Cumulative readings are modeled as scopeless relations: Even when the plural entities A and B are introduced by scope-taking elements such as indefinites, none of them takes scope over the other. For example, the cumulative reading of Sentence (67a) is modeled as follows:

- (68) $\exists e[*use(e) \wedge *dutch.firm(*agent(e)) \wedge |*agent(e)| = 600 \wedge$
 $*american.computer(*theme(e)) \wedge |*theme(e)| = 5000]$

This representation is equivalent to the representation in (69) given the following assumptions: verbs and thematic roles are closed under sum (see Sections 2.7.2 and 2.5.1), and *use* is distributive down to atoms on both its agent and theme arguments (see Section 4.5). While the representation in (69) is probably more readable, the one in (68) is easier to derive compositionally (see Section 2.10) because it keeps the two arguments neatly apart. This observation goes back to Schein (1986) and Krifka (1989b), and is developed in detail in Krifka (1999) and Landman (2000).

- (69) $\exists x \exists y[*dutch.firm(x) \wedge |x| = 600 \wedge *american.computer(y) \wedge |y| = 5000$
 $\wedge \forall x' \leq_{Atom} x \exists y' \leq_{Atom} y \exists e[use(e) \wedge agent(e) = x' \wedge theme(e) = y']$
 $\wedge \forall y'' \leq_{Atom} y \exists x'' \leq_{Atom} x \exists e[use(e) \wedge agent(e) = x'' \wedge theme(e) = y'']]$

I assume that cumulative readings are a separate kind of interpretation, and must be distinguished from collective readings (see Landman 1996, 2000).

Typical examples of collective readings involve a noun phrase in agent position, and entail that the group to which the noun phrase refers bears collective responsibility for the event, above and beyond the individual responsibilities of its members. Typical examples of cumulative readings do not give rise to this entailment.

For example, in (70) the collective reading is more prominent than the cumulative reading. The collective reading of this sentence does not entail that each cowboy sent an emissary to some Indian, nor that each Indian was sent an emissary. In addition, the cowboys bear collective responsibility for the action of sending an emissary.

(70) The cowboys sent an emissary to the Indians.

By contrast, example (71) only has a cumulative reading, not a collective reading, firstly because it does entail that each man in the room is married to a girl across the hall and vice versa, and secondly because there is no sense in which the men have collective responsibility for being married to the girls above and beyond their individual responsibilities.

(71) The men in the room are married to the girls across the hall. (Kroch 1974)

Collective readings also occur when there is only one noun phrase in the sentence:

(72) Three boys carried a piano upstairs.

Following Landman (2000) and references therein, I assume that this sentence has the following readings and representations:

(73) Potentially-different-pianos distributive reading:

$$\exists x[*\text{boy}(x) \wedge |x| = 3 \wedge \forall y. y \leq_{\text{Atom}} x \rightarrow \\ \exists e[*\text{carry.upstairs}(e) \wedge *\text{agent}(e) = y \wedge \text{piano}(*\text{theme}(e))]]$$

(Three boys each carried a piano upstairs.)

(74) Same-piano distributive reading:

$$\exists e[*\text{carry.upstairs}(e) \wedge \\ *\text{boy}(*\text{agent}(e)) \wedge |*\text{agent}(e)| = 3 \wedge \text{piano}(*\text{theme}(e))]$$

(There is a piano that three boys each carried upstairs.)

(75) Collective reading:

$$\exists e[*\text{carry.upstairs}(e) \wedge \\ \exists x[*\text{boy}(x) \wedge |x| = 3 \wedge *\text{agent}(e) = \uparrow(x) \wedge \text{piano}(*\text{theme}(e))]]$$

(There is a piano that three boys, as a group, carried upstairs.)

These representations rely on the assumption that *carry.upstairs* is distributive on at least its agent predicate. That is, whenever a sum of individuals is the agent of a carrying-upstairs event *e*, its mereological parts are the agents of carrying-upstairs events that are parts of *e*. The notion *distributive predicate* is taken up again in Section 4.5.

The collective reading blocks this entailment by relating the piano to an atom that represents the boys as a group. Following Landman (2000), I represent this reading

by shifting the sum individual x that represents three boys to a group individual $\uparrow(x)$ that represents these three boys, taken as a group (see Section 2.3.1). Since this group is taken to be a mereological atom, the relation between a group and its members is distinct from the mereological relation between a proper sum and its parts. Following Landman, I call atoms “impure” when they are formed through the group formation operator. As described in Section 2.3.1, I assume that atomicity checks can distinguish between pure and impure atoms. I assume that the atoms in the denotation of singular count nouns are all pure atoms (see Section 2.6.1).

To derive the readings above, I assume that definites and indefinites are ambiguous between a sum and a group interpretation. Some sample translations of verb phrases with definites and indefinites are given in (76) (see Section 2.10 for the compositional process that underlies them). Interpreting plural definite descriptions using \oplus is actually a simplification of Landman’s account, as it fails to rule out *the two cats* in a model that contains three cats. This simplification is harmless for the purpose of this book.

- (76) a. $\llbracket \text{a boy} \rrbracket = \llbracket \text{boy} \rrbracket = \lambda x[\text{boy}(x)]$
 b. $\llbracket \text{boys} \rrbracket = \lambda x[*\text{boy}(x)]$
 c. $\llbracket \text{three boys}_{\text{sum}} \rrbracket = \lambda x[*\text{boy}(x)] \wedge |x| = 3$
 d. $\llbracket \text{three boys}_{\text{group}} \rrbracket = \lambda x \exists y[x = \uparrow(y) \wedge *\text{boy}(y) \wedge |y| = 3]$
 e. $\llbracket \text{the boys}_{\text{sum}} \rrbracket = \oplus \text{boy}$
 f. $\llbracket \text{the boys}_{\text{group}} \rrbracket = \uparrow(\oplus \text{boy})$

As these translations illustrate, I assume that numerals have an *exactly* interpretation. As discussed in Section 2.5.5, this assumption is not crucial because in this book, I only consider upward-entailing quantifiers anyway. The generalization to non-upward-entailing quantifiers is compatible with the approach presented here but requires additional technical devices such as maximization operators (Landman 2000, Krifka 1999, Brasoveanu 2013).

Section 2.4.6 gives a compositional derivation of a numeral noun phrase on its sum interpretation. I assume that the other types of noun phrases have similar derivations. For example, the group interpretation can be obtained through an optional type-shifting mechanism. The details do not matter. See Landman (2000) for a concrete proposal.

2.9 Prepositional phrases

There is essentially only one prepositional phrase that plays a major role in this book, namely the directional prepositional phrase (*all the way*) *to the store* (see Chapter 6). I assume a semantic representation in the style of Zwarts (2006) that involves a continuous directed spatial interval or path whose end is (the location of) the store.

- (77) $\llbracket (\text{all the way}) \text{ to the store} \rrbracket$
 $= \lambda V_{\langle v, t \rangle} \lambda e [V(e) \wedge \text{continuous}(\sigma(e)) \wedge \text{end}(\sigma(e)) = \text{the.store}]$

See also Hinrichs (1985), Krifka (1998), Zwarts (2005, 2006) for the treatment of paths in aspectual composition, and Section 2.4.4 for the notion of spatial interval.

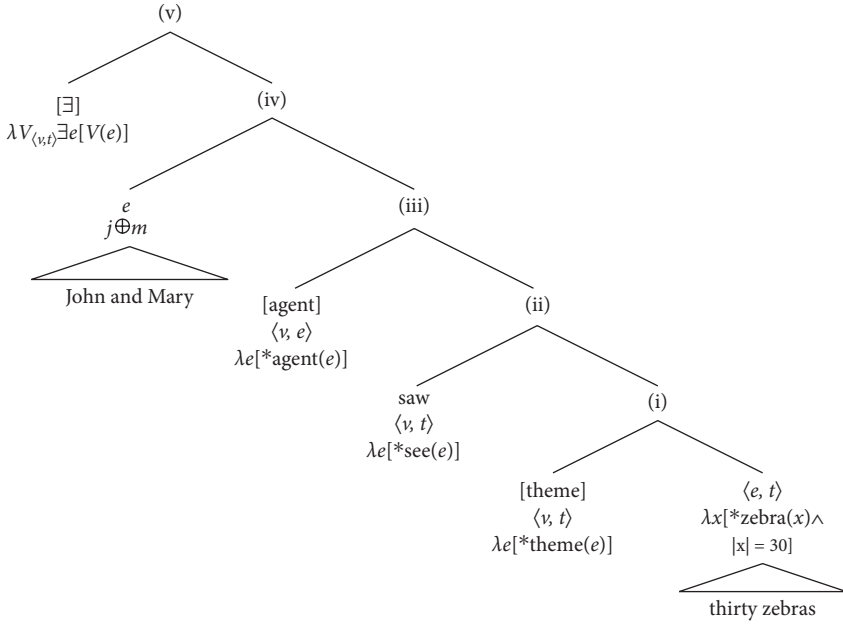
2.10 The compositional process

Following Parsons (1990) and Krifka (1992), I assume that verbs and verbal projections denote sets of events, type $\langle v, t \rangle$ (Section 2.7.1). For noun phrases, as described in Section 2.8, I follow the standard assumption that they can have individual type e , predicative type $\langle e, t \rangle$ or quantificational type $\langle et, t \rangle$, though I will not make use of the quantificational type (see Section 9.5 for my treatment of quantificational noun phrases headed by *every* and *each*). I assume that the application of a predicative or referential noun phrase to a verbal projection amounts to intersecting two sets of events. This idea goes back at least to Carlson (1984). For example, the interpretation of *John loves Mary* amounts to intersecting the set of events whose agent is John, the set of loving events, and the set of events whose theme is Mary. At the end of the derivation, a sentence mood operator applies. Since I only consider declarative sentences, I make the standard assumption that this operator only binds the event variable with an existential quantifier. I call this operator existential closure or simply $[\exists]$.

Following Landman (2000), I assume that referential and predicative noun phrases can be interpreted in situ or by quantifier raising (see Section 2.8). The latter option can also account for distributive readings of these noun phrases, given certain assumptions about the quantifier-raising operation. Since the relevant readings do not play a role in this book, I refer the reader to Landman (2000) for details.

Implementing Carlson's idea requires shifting the types of verbal projections and noun phrases. There does not seem to be a best practice of how to resolve this kind of type mismatch; instead, authors follow their own conventions. For example, Kratzer (1996), who only uses the thematic role corresponding to the external argument, assumes an operation of "event identification" that connects the thematic role with the verb phrase, while Landman (2000) assumes that the compositional process can lift the type of verbal projections so that they combine with noun phrases through function application.

I assume, similarly to Landman, that the type mismatch is resolved by type shifters. My setup is different from his because he assumes neither that verbs denote sets of events, nor that thematic roles are introduced by separate lexical entries. I assume that the following type shifters can apply freely to thematic roles, verbal projections, and noun phrases:



(i) Predicative type shifter, NP first

$$\lambda V_{\langle v, t \rangle} \lambda e [* \text{zebra}(*\text{theme}(e)) \wedge |* \text{theme}(e)| = 30]$$

(ii) Function application

$$\lambda e [* \text{see}(e) \wedge * \text{zebra}(*\text{theme}(e)) \wedge |* \text{theme}(e)| = 30]$$

(iii) Referential type shifter, VP first

$$\lambda x \lambda e [* \text{agent}(e) = x \wedge * \text{zebra}(*\text{theme}(e)) \wedge |* \text{theme}(e)| = 30]$$

(iv) Function application

$$\lambda e [* \text{see}(e) \wedge * \text{agent}(e) = j \oplus m \wedge * \text{zebra}(*\text{theme}(e)) \wedge |* \text{theme}(e)| = 30]$$

(v) Existential closure

$$\exists e [* \text{see}(e) \wedge * \text{agent}(e) = j \oplus m \wedge * \text{zebra}(*\text{theme}(e)) \wedge |* \text{theme}(e)| = 30]$$

FIG. 2.6 The compositional process.

(78) Predicative type shifters:

- a. VP, then NP: $\lambda\theta_{\langle v, e \rangle} \lambda V_{\langle v, t \rangle} \lambda P_{\langle e, t \rangle} \lambda e [V(e) \wedge P(\theta(e))]$
- b. NP, then VP: $\lambda\theta_{\langle v, e \rangle} \lambda P_{\langle e, t \rangle} \lambda V_{\langle v, t \rangle} \lambda e [V(e) \wedge P(\theta(e))]$

(79) Referential type shifters:

- a. VP, then NP: $\lambda\theta_{\langle v, e \rangle} \lambda V_{\langle v, t \rangle} \lambda x_e \lambda e [V(e) \wedge \theta(e) = x]$
- b. NP, then VP: $\lambda\theta_{\langle v, e \rangle} \lambda x_e \lambda V_{\langle v, t \rangle} \lambda e [V(e) \wedge \theta(e) = x]$

For concreteness, I assume that these type shifters always apply first to the thematic role, and then in any order to the verbal projection and to the noun phrase. This flexible order allows me to remain noncommittal about whether noun phrases c-command thematic role heads or are c-commanded by them. The former assumption is generally made for the little-*v* head, which I represent as [agent], while the latter assumption makes thematic role heads analogous to silent prepositions. In order to simplify the compositional process, I will assume that both options are possible for the [agent] head.

I also assume that event predicates can combine with other event predicates by a generalized form of intersection, similarly to the predicate modification rule in Heim & Kratzer (1998) and to the event identification rule introduced by Kratzer (1996). The idea that verbs and their arguments are combined by intersection is also argued for in Carlson (1984) and is elevated to a general principle, conjunctivism, in Pietroski (2005, 2006).

The compositional process is illustrated in Figure 2.6. The sentence used here is *John and Mary saw thirty zebras*. Based on our background assumptions, the LF generates the cumulative reading of this sentence: there is a sum of seeing events whose agents sum up to John and Mary and whose themes sum up to thirty zebras. This reading does not entail that each zebra was seen by both John and Mary, nor that the zebras were seen simultaneously. I return to this point in Chapter 4.

The same LF could also model the sentence *John and Mary looked at thirty zebras* if we assume that *see* and *look* have the same lexical entries and that *at* is the overt counterpart of [theme]. This is the main motivation for the syntactic position of [theme].

My assumption that the type of verbs and their projections is $\langle v, t \rangle$ and that they denote predicates of events is standard, but not universally accepted. I myself have argued elsewhere to the contrary that the type of verbs and verbal projections should be taken to be $\langle vt, t \rangle$ if we want to account for the interaction of verbs with quantification and other scope-taking phenomena (Champollion 2011a, 2015d). Here I stick with the more standard $\langle v, t \rangle$ assumption, partly in order to make sure the system remains compatible with the majority of existing theories and partly because the lower type is sufficient for present purposes. The relation between Champollion (2015d) and the present proposal is explored in more detail in Schwarzschild (2014) and Champollion (2014).

The cast of characters

3.1 Introduction

Three constructions are at the center of this book: *for*-adverbials, pseudopartitives, and constructions with *each*.

- (1) *For*-adverbials:
 - a. John **ran** for five minutes. *atelic*
 - b. *John **ran to the store** for five minutes. **telic*
- (2) Pseudopartitives:
 - a. thirty pounds of **books** *plural*
 - b. thirty liters of **water** *mass*
 - c. *thirty pounds of **book** **singular*
- (3) Constructions with *each*:
 - a. The boys **each walked**. *distributive*
 - b. *The boys **each met**. **collective*

As mentioned in Chapter 1, these constructions all have a constituent which is subject to a certain restriction. The predicate modified by a *for*-adverbial must be atelic, the substance noun of a pseudopartitive must be either mass or plural, and the verb phrase of an *each*-construction must be distributive. In examples (1) through (3) I have marked these constituents in bold.

This chapter provides a scaffold on which the theory in the rest of the book is built. Based on the foundations laid out in Chapter 2, I present what I take to be a plausible baseline theory for the syntax and semantics of these constructions and their constituents, keeping things symmetric across domains as much as seems reasonable so that the parallels which I draw in subsequent chapters are not obscured more than necessary. More refined theories might require giving up the symmetries I assume here. Doing so should not affect the insights underlying the theory presented in the subsequent chapters, but it might make it necessary to reformulate the technical implementation in a less conspicuous way.

I discuss pseudopartitives in Section 3.2, *for*-adverbials in Section 3.3, and adverbial-*each* constructions in Section 3.4.

3.2 Pseudopartitives

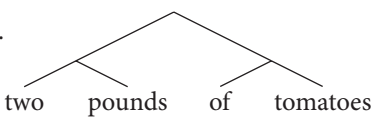
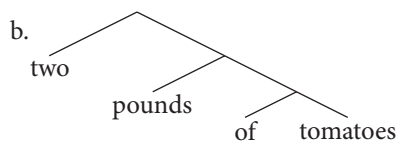
Pseudopartitives, also called measure constructions, are noun phrases that involve reference to an amount of some substance. Both the amount and the substance involved are specified with a noun; in English, the nouns are separated by the word *of*:

- (4) three liters of drinkable water

The term *pseudopartitive* was introduced by Selkirk (1977) to distinguish this construction from true partitives such as *three liters of the water*. I refer to the noun that comes to the right of *of* as the *substance noun*, and to the substance noun together with any of its modifiers the *substance nominal*. The noun *liters* in this example is a *measure noun* (see Section 2.6.3). The substance noun is *water*, and the substance nominal is *drinkable water*. As explained below, I treat *three liters* as a constituent, namely a *measure phrase*.

Pseudopartitives with measure nouns are called *measure pseudopartitives*. Pseudopartitives can also be formed in other ways, for example with container nouns (*three glasses of wine*) or classifier nouns (*three heads of cattle*). For a discussion of the syntactic and semantic properties of the different kinds of pseudopartitives, see Keizer (2007: ch. 6). Group nouns like *committee* can head a construction that is at least superficially similar to pseudopartitives (*a committee of women*). As discussed in Section 2.6.4, I do not consider this construction a pseudopartitive.

The claims in this book only concern the measure reading of pseudopartitives. Pseudopartitives with measure nouns only have a measure reading; pseudopartitives with container nouns are ambiguous between a measure reading and an individuating reading (see Section 2.6.3). As the term *measure phrase* suggests, I assume that the measure noun of a measure pseudopartitive forms a constituent with the determiner that precedes it, as in (5a). An alternative to this analysis would involve the right-branching structure (5b). I adopt (5a) to reflect the semantic parallel between pseudopartitives and other distributive constructions (see Table 4.3 in Chapter 4). However, it would not be difficult to reformulate the claims in this book in a way that is consistent with structure (5b). My compositional implementation in Section 4.7 places the essential machinery into the lexical entry of *of*. The only relevant difference between the two structures is whether *of* combines with *two* and *pounds* at once or one at a time. It is an easy technical exercise to rewrite the lexical entry of *of* to mirror this difference.

- (5) a.  b. 

The coordination test in (6) and (7) shows that my syntactic assumption is at least initially plausible.⁸ While the string *two pounds* can be coordinated without problems in (6), many people reject sentences in which the string *pounds of tomatoes* is part of a coordination structure in (7).

- (6) John bought two pounds and two ounces of tomatoes.
 (7) *John bought two pounds of tomatoes and grams of saffron.

More arguments in favor of the symmetric structure are found in Gawron (2002). However, the choice between the two structures is by no means settled. The symmetric structure is also adopted by Akmajian & Lehrer (1976), Guéron (1979), Gawron (2002), and Schwarzschild (2002, 2006), while the right-branching structure is adopted by Stickney (2008), Chierchia (2008), Bale (2009), and Scontras (2014), among others. Landman (2004) and Rothstein (2009) make the plausible claim that the symmetric structure corresponds to the measure reading and the right-branching structure to the individuating reading.

I represent the meaning of a measure pseudopartitive as follows:

- (8) three liters of water
 $\lambda x[\text{water}(x) \wedge \text{liters}(\text{volume}(x)) = 3]$

I write $\text{liters}(\text{volume}(x)) = 3$ and not simply $\text{liters}(x) = 3$ because I assume that there is a layer of degrees that mediates between substances and natural numbers (Section 2.4.5). I use the term *measure function* for functions from entities to degrees, and the term *unit function* for functions from degrees to numbers (see Sections 2.5.3 and 2.5.4).

To derive the representation in (8) compositionally, I assume the lexical entries and the LF in Figure 3.1. I represent the source of the measure function *volume* as a silent lexical item, which I write [volume]. This is one possible way to express the fact that the overt part of a pseudopartitive does not always uniquely specify this measure function (see Section 2.5.4). It would also be possible to introduce the measure function as a free variable. As far as I can tell, this choice does not interact with any other assumptions. The position of [volume] in the tree does not matter either.

The arguments of the translation of *of* are represented with the letters S, M, K, and b. The uppercase letters can be taken as standing for predicates of arbitrary predicative type here. In Section 4.4, they resurface as mnemonics for the terms Share, Map, and Key, which were introduced in Chapter 1. The letter b is a variable over entities of any basic type. The type of *of* is polymorphic because I assume that events rather than ordinary objects, and intervals rather than degrees, can also be used. The Greek letters α and β stand for arbitrary types. In this particular LF, α is resolved to *e* and β to *d*.

⁸ I thank Alan Bale for discussing this issue with me.

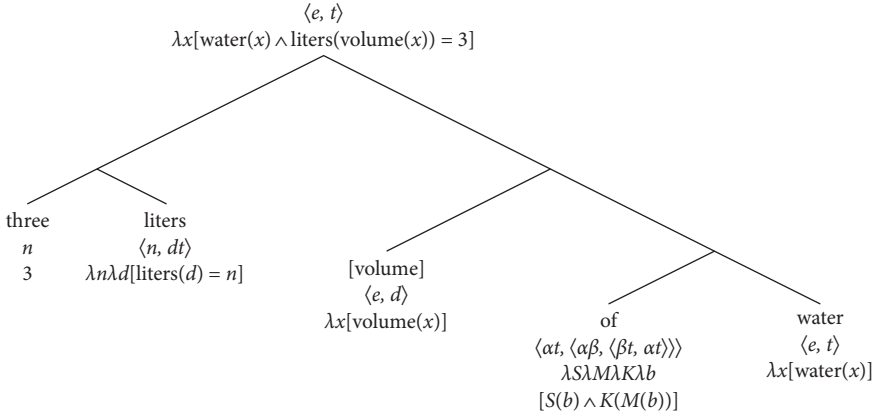


FIG. 3.1 Skeletal LF of an ordinary pseudopartitive.

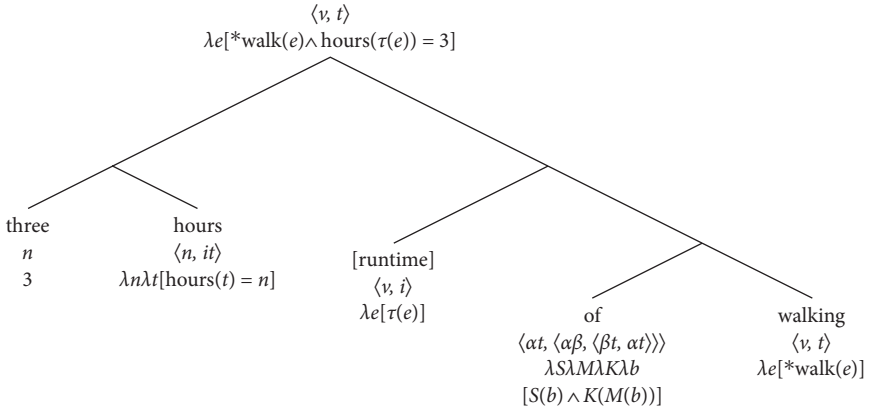


FIG. 3.2 Skeletal LF of an event pseudopartitive.

But this may be different in other cases. For example, I assume that a pseudopartitive like *three hours of walking* denotes a set of events rather than individuals, and that the domain of the unit function introduced by the word *hours* contains intervals rather than degrees (see Section 2.5.4).

Figure 3.2 shows an LF for *three hours of walking*. The star in front of the translation of *walk* is a reminder that the denotations of verbs are assumed to be their own algebraic closures; see Section 2.7.2.

This type of representation is an incomplete skeleton. Alone, it does not provide any explanation for the constraints on pseudopartitives described in Chapters 1 and 7, because it does not contain any constraints itself. For example, there is no explanation why **three pounds of book* cannot be interpreted as follows:

- (9) three pounds of book
 $\lambda x[\text{book}(x) \wedge \text{pounds}(\text{weight}(x)) = 3]$

A constraint that rules out this pseudopartitive, along with the other unacceptable examples of the constructions discussed in Chapter 1, is introduced in Chapter 4. It is further developed in Chapter 7.

3.3 For-adverbials

For-adverbials have also been called *measure adverbials*, *aspectual adverbials*, and (when they are temporal rather than spatial) *durative adverbials*. Based on examples like (10), they have been used as a diagnostic of atelicity since at least Verkuyl (1972). In fact, Verkuyl (1989) considers them the most reliable indicator of atelicity.

- (10) a. John ran { for five minutes / for three hours / for miles.}
 b. *John ran to the store { for five minutes / for three hours / for miles.}

As in the case of the count/mass distinction (see Section 2.6), the telic/atelic distinction shows a certain amount of elasticity. To some extent, *run to the store* may be reinterpreted or “coerced” as an atelic predicate with the same meaning as *run towards the store*. I have nothing to say about the phenomenon of aspectual coercion. I also exclude from consideration the so-called result-state interpretation of *for*-adverbials (Piñón 1999b), which is discussed by Dowty (1979) under the name of “internal reading.” Example (11) is ambiguous between interpretations (12a) and (12b). Dowty attributes this observation to Binnick (1969).

- (11) The Sheriff of Nottingham jailed Robin Hood for four years.
- (12) a. The Sheriff of Nottingham spent four years bringing it about that Robin Hood was in jail.
 b. The Sheriff of Nottingham brought it about that for four years Robin Hood was in jail.

The result state related interpretation must be controlled for when we use *for*-adverbials as an atelicity diagnostic, because it allows *for*-adverbials to combine with telic and punctual predicates (*John opened the window for five minutes*). On this interpretation, which corresponds to (12b), *for*-adverbials are compatible with accomplishments (e.g. *John opened the window for five minutes*), and therefore do not diagnose telicity. Many languages, such as German, have different words for the standard and the result-state interpretation of *for*-adverbials (Piñón 1999b).

For-adverbials stand in near-complementary distribution with *in*-adverbials, which reject atelic predicates and accept telic predicates. As Section 4.3.1 discusses, the entailment pattern in (13) shows that sentences with *for*-adverbials, but not those with *in*-adverbials, are distributive constructions:

- (13) a. John ran for five minutes.
 \Rightarrow John ran for four minutes.
 \Rightarrow John ran for three minutes.
 b. John ran to the store in five minutes.
 \nRightarrow John ran to the store in four minutes.

I do not discuss *in*-adverbials in this book.

It has often been observed that *for*-adverbials can tolerate varying amounts of discontinuity. Even when we keep the verb and adverbial constant, this amount can vary considerably from one case to another, subject to pragmatic constraints. As Stefan Hinterwimmer (p.c.) points out, this is reminiscent of the behavior of generic and habitual statements. However, the phenomenon occurs even when the *for*-adverbials are used to make clearly episodic statements, as can be observed in the following pair of examples (Barbara Partee, p.c. in Vlach 1993). Sentence (14a) requires Mary to sleep almost continuously; sentence (14b) is compatible with an ordinary sleeping pattern of about eight hours a day.

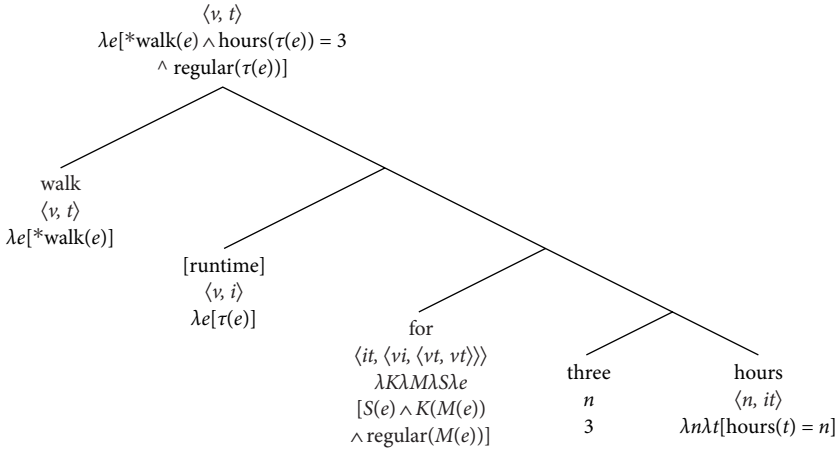
- (14) a. Mary slept for a week.
 b. Mary slept in the attic for a week.

Since *sleep* and *sleep in the attic* are both atelic, this constraint needs to be seen as separate from the constraint against telic predicates. It ultimately needs to be addressed by any theory of *for*-adverbials, and there are a few that do so (Piñón 1999b, Landman & Rothstein 2012a, 2012b). I will not discuss these theories or model the regularity constraint in this book; see Chapter 8 here, Deo & Piñango (2011), and Champollion (2013) for related discussion. In my presentations of theories of *for*-adverbials, I will simply use the shorthand predicate *regular* as a stand-in for a more worked-out theory of this constraint.

Regarding the syntax of *for*-adverbials, there is no consensus on whether they attach below or above the subject. This issue becomes relevant in connection with the interaction of *for*-adverbials and the Perfect. An overview of the relevant issues and literature is found in Rathert (2004). The issue hinges on whether sentences like (15a) and (15b) are synonymous:

- (15) a. John has been in Boston for four years.
 b. For four years, John has been in Boston.

Dowty (1979) reports that (15a) does not specify whether John is still in Boston, but he judges (15b) to entail that he is still in Boston. Researchers who do assume an ambiguity often attribute it to different attachment sites: in (15a), the *for*-adverbial can attach only above the subject; in (15b), it can also attach below the subject. However, the accuracy of Dowty's judgment has been challenged (Abusch & Rooth 1990, Rathert 2004).

FIG. 3.3 Skeletal LF of a *for*-adverbial.

In any case, the premise of such arguments is that *for*-adverbials always take semantic scope where they take syntactic scope. Given this premise, another argument that the higher attachment site is available comes from the contribution of the subject to the aspectual properties of the sentence. Verb phrases that are otherwise unacceptable with a *for*-adverbial can be rescued by an unbounded subject (Verkuyl 1972):

- (16) a. For years, { water / *a liter of water } came in from the rock.
 b. For hours, { people / *this person } walked out of the house.
 c. { Patients here / *These two patients } died of jaundice for months.

Conversely, the premise can also be used to argue for a possible attachment site below the subject by using examples of verb phrase coordination:

- (17) Yesterday one of my friends worked for eight hours, had dinner, and then slept for five hours.

In this sentence, the two *for*-adverbials semantically, and therefore arguably syntactically, modify the individual verb phrases rather than the entire sentence.

Figure 3.3 shows my LF for a verb phrase to which a *for*-adverbial has attached. As in the case of pseudopartitives, this is an incomplete account that serves as a canvas for what follows. Nothing in the entry of the *for*-adverbial presented here rules out telic predicates.

As Figures 3.2 and 3.3 show, I assume that there is no semantic difference between a verb phrase modified by a *for*-adverbial like *run for three hours* and an event pseudopartitive like *three hours of running*.⁹ Of course, the two phrases do not have

⁹ A possible exception is the regularity constraint, which I set aside in the case of pseudopartitives. I do not take a stance on whether it applies to them. As discussed in Section 2.5.4, events and their time values

the same distribution. However, since I assume that all verbal projections involve reference to events, I have not used polymorphic types in my translation of the word *for*.

The parallel between *for*-adverbials and event-based pseudopartitives that I assume here is analogous to the well-known parallel between sentences and action nominalizations. On the assumption that both involve reference to events, they can be given identical semantic representations:

- (18) a. The Romans destroyed a city.
 b. the Romans' destruction of a city
 c. $\lambda e[*\text{destroy}(e) \wedge *\text{agent}(e) = \text{the.romans} \wedge \text{city}(\text{theme}(e))]$

3.4 Adverbial-*each* constructions

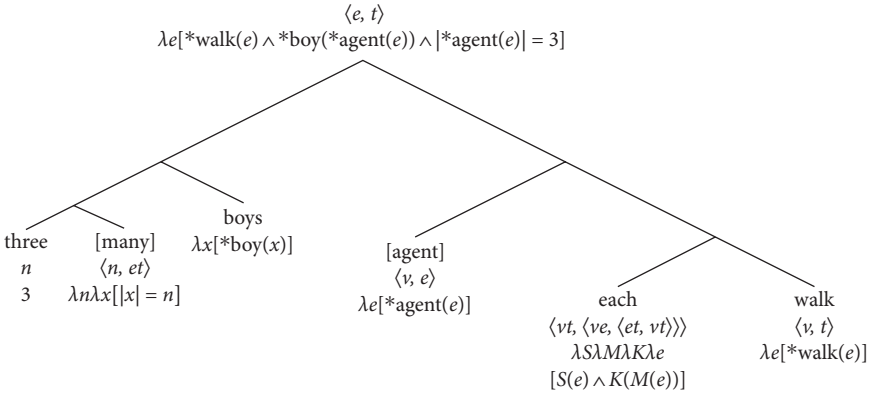
The distributive item *each* can occur in different positions: as an adnominal modifier (19a), as an adverbial modifier (19b), and as a determiner-quantifier (19c).

- (19) a. **Adnominal:** Two men have carried three suitcases *each*.
 b. **Adverbial:** Two men have *each* carried three suitcases.
 c. **Determiner:** *Each* man has carried three suitcases.

In this chapter and in Chapter 4, I focus on the adverbial modifier use of *each*, and I refer to the type of sentences that contain it as adverbial-*each* constructions. I use *each* mainly for the purpose of building a bridge between distributivity and the other phenomena considered here. To this end, I show how to model its distributivity effect in an event-semantic framework. In order to draw out the parallels between *each*, *for*, and *of*, I adopt an analysis that makes them maximally similar to each other. Chapter 9 shifts to a slightly different view. Since it focuses on the parallels between overt distributive items and the covert distributivity operators that are the topic of Chapter 8, it treats distributive items as overt realizations of distributivity operators (Link 1991b). The difference between the two perspectives is not crucial because both analyses are ultimately grounded in stratified reference. See Section 9.3 for more discussion.

I assume that adverbial *each* modifies a verb phrase. As discussed in Section 2.10, I assume that thematic role heads occur between noun phrases and the verbal projections with which they combine. For the agent role head, this assumption is widespread in the syntactic literature. The usual name for this head is *v* or “little *v*.” Since I omit category labels in syntactic representations, I represent it instead as [agent]. The square

are mediated by a layer of potentially discontinuous intervals, while substances and their volume values are mediated by a layer of degrees to which the notion of discontinuity does not apply. For this reason, in the absence of a theory of regularity, it is not clear what exactly the regularity constraint would amount to in the case of a substance pseudopartitive and whether it could ever be violated.

FIG. 3.4 Skeletal LF of an adverbial-*each* construction.

brackets are a reminder of the parallelism I assume between thematic roles, trace functions like [runtime], and measure functions like [height]. I assume that thematic roles are functions as well (see Section 2.5.1).

Like *of* in pseudopartitives and *for* in *for*-adverbials, adverbial *each* can only be associated with predicative noun phrases (**every boy each walked*). As in the cases of *of* and *for*, I write this fact directly into the lexical entry of *each*. To accommodate the fact that referential noun phrases such as *the boys* or *John and Mary* are also compatible with *each*, I assume that any noun phrase *N* whose translation is of type *e* can be shifted to a noun phrase of type $\langle e, t \rangle$ which denotes the singleton set $\lambda x[x = \llbracket N \rrbracket]$. This is the *Quine* operator of Partee & Rooth (1983). For a similar assumption regarding the definite article, see Winter (2001: 153).¹⁰

Figure 3.4 shows my LF for the sentence *three boys each walked* before existential closure applies (Section 2.10). For convenience, I have not decomposed *boys* into *boy* plus plural. See Section 2.6.2 for my assumptions on the meaning of plural count nouns. As in the previous sections, this is an incomplete LF which does not yet rule out collective predicates like *met* and which does not make sure that predicates like *build a raft* are interpreted distributively. These possibilities will be ruled out in Chapter 4, when a stratified-reference constraint is added to this LF. Chapter 8 will recast this constraint as a distributivity operator and Chapter 9 will equate the meaning of *each* with this operator.

¹⁰ The *Quine* operator could also be used to cut down by half the number of type shifters in Section 2.10, at the expense of producing slightly less readable logical translations.

The theory

4.1 Introduction

This chapter motivates and develops strata theory. I have already introduced its main idea in Chapter 1. Here, I develop it in full formal detail against the backdrop of the ontological and semantic assumptions presented in Chapter 2 and the LFs presented in Chapter 3.

As described in the introduction, strata theory capitalizes on cross-categorical parallels between atelic aspect, mass reference, plural reference, and distributivity. Many authors have made analogies between some of these properties and put them to use for drawing semantic generalizations across constituents that have these properties. For example, the parallel between mass and plural reference is exploited in Link (1983) and the one between mass and atelic reference in Bach (1986). Rather than modeling these parallels through syntactic features, these authors take the important step of relating them to semantic properties of the denotations of the constituents in question. In a mereological framework, these properties can be formalized through higher-order properties such as cumulative and divisive reference. This is also the strategy I use here.

As this chapter will show, the higher-order properties that Link and Bach used to characterize mass, plural, and atelic reference do not adequately characterize the restrictions that distributive constructions impose on their constituents. For this reason, I introduce a new higher-order property called *stratified reference*. This property has parameters which provide additional degrees of freedom compared to better-known properties like divisive and cumulative reference. These parameters make it possible to use stratified reference to express similarities and differences between distributive constructions in a uniform way.

After an informal overview of different concepts associated with the word *distributivity*, Section 4.2 proposes an operational definition of distributivity when understood as a property of predicates. Section 4.3 introduces the notion of a distributive construction—a lexicosyntactic configuration that imposes an obligatory distributive relation between two of its constituents—and motivates treating the constructions introduced in Chapter 3 as distributive constructions. Section 4.4 introduces the terms *Key*, *Share*, and *Map* as construction-neutral ways to refer to the components of a

distributive relation. After introducing the concept of meaning postulates, Section 4.5 reviews and refines previous theories of the constraints that adverbial-*each* constructions, *for*-adverbials, and pseudopartitives impose on their constituents. These refinements result in a convergence that Section 4.6 exploits to unify these theories into a single constraint, which is expressed using the notion of stratified reference. Section 4.7 uses the scaffolding provided in Chapter 3 to implement this constraint compositionally, and Section 4.8 concludes.

4.2 What is distributivity?

The use of the word *distributivity* generally indicates the application of a predicate to the members or subsets of a set or group, or to the parts of an entity. This application is diagnosed by the presence of certain entailments I call *distributive entailments*. While the word *distributivity* and derived terms are widespread in the semantic literature, there are no standard definitions, and a number of related concepts can be distinguished for which the word is used. Distributivity can be seen as property of quantifiers (Section 4.2.1), a relation between two constituents (Section 4.2.2), a property of predicates (Section 4.2.3), or, as I propose in Section 4.3, a property of constructions. All these concepts are useful, but it is important to delimit them from each other. Partly building on an overview by Tsoulas & Zweig (2009), this section reviews existing notions of distributivity and provides an overview of the relevant results from the semantic literature.

4.2.1 Quantificational distributivity

We speak of quantificational distributivity in connection with quantificational noun phrases headed by determiners like *every* or *each* (e.g. Scha 1981). The denotations of these noun phrases involve the application of the verbal predicate to each member of their witness set. This is probably the least controversial type of distributivity. I will set it aside until Section 9.5, where I propose an analysis of the two determiners *every* and *each*, and Chapter 10, where I propose an analysis of *all* in its adnominal use. There is no consensus on whether adnominal *all* should be classified as a distributive quantifier, because there is a class of predicates with which it gives rise to collective readings (I will discuss them in Sections 4.2.3 and 10.4 under the name “gather-type predicates”). Chapter 10 presents arguments for treating *all* as a distributive quantifier, and draws a parallel between its behavior and other distributive items.

4.2.2 Relational distributivity

When two different constituents contribute to the content of a distributive entailment, they are said to stand in a distributive relation. I refer to the presence of a distributive relation generally as *relational distributivity*. In standard examples like

(1a), the constituents involved in a distributive relation are a subject and a verb phrase, but this need not be so. A theory of distributivity that relies heavily on the concept of a distributive relation is developed in Choe (1987). By contrast, I do not assume that distributive relations are reified as syntactic or semantic links.

- (1) a. Al and Bill each ate a pizza.
b. Al and Bill ate a pizza.

Sentence (1a) has the distributive entailments that Al ate a pizza, and that Bill ate a pizza. A distributive relation can be obligatory or optional. For example, sentence (1b) leaves open whether one or two pizzas were eaten. In (1a), the relation between the subject and the verb phrase is obligatorily distributive. In (1b), it is optionally distributive.

A distributive relation can be indicated by distributive items such as adverbial *each* in (1a) and related items in other languages (see Chapter 9 for a cross-linguistic survey). In the absence of overt items, certain sentences such as (1b) can exhibit a distributive relation optionally (see Chapter 8 for an analysis of covert optional distributivity).

4.2.3 *Predicative distributivity*

Distributivity understood as a property of predicates is generally set in opposition to collectivity. These notions are based on the behavior of predicates when they occur with plural definites, noun phrases headed by *every*, and noun phrases coordinated by *and*. Predicates such as *smile* or *sing* lead to (near-)equivalent sentences when these different kinds of arguments are used, as in (2). These predicates are classified as distributive. The class of collective predicates is formed by those predicates for which this pattern breaks down because the combination with *every* and with singular proper names leads to a category mistake, as in (3).

(2) **Distributive predicates**

- a. The five children smiled. \Leftrightarrow Every one of the five children smiled.
- b. John and Mary smiled. \Leftrightarrow John smiled and Mary smiled.

(3) **Collective predicates**

- a. The children gathered. \nLeftrightarrow *Every child gathered.
- b. John and Mary met. \nLeftrightarrow *John met and Mary met.

This distinction between distributive and collective predicates has been criticized by Winter (2001, 2002) as not very useful and hard to justify. Winter notes that the patterns in (2) and (3) are only valid if one abstracts away from a number of factors: conventionalized coordinations, nonmaximality effects on definite plurals, and effects related to group nouns. Winter's concern about conventionalized coordinations like *Simon and Garfunkel*, if I understand him correctly, is that they do not always give rise to entailments like (2b), as shown in (4), which Winter bases on a similar example

he attributes to Fred Landman (p.c.). The point is that only Garfunkel may be singing while Simon is playing the guitar. The biconditional in (2a) is only valid to the extent that the referent of the definite plural includes every member of its complement noun, but this is not the case if the definite plural has a nonmaximal interpretation, as in (5) (see Križ 2016). The point here is that in a typical press conference, only a few reporters will get to ask questions. Furthermore, the test in (3) is only reliable as long as its nouns are not replaced by group nouns and noun phrases like *committee* and *Committee A*, as shown in (6) (see Barker 1992, de Vries 2015). With such noun phrases, there is a distributive reading on which the entailment properties of these sentences are similar to those of distributive predicates that were illustrated in (2). (I draw the arrow in only one direction because the sentences before it also have a collective reading, on which they do not entail the sentences after it.)

- (4) Simon and Garfunkel are performing in Central Park.
 \nRightarrow Simon is performing in Central Park.
- (5) At the end of the press conference, the reporters asked the president questions.
 \nRightarrow Every reporter asked the president a question. (Dowty 1987)
- (6) a. The ten committees gathered.
 \Leftarrow Every one of the ten committees gathered.
 b. Committee A and Committee B met.
 \Leftarrow Committee A met and Committee B met.

Winter concludes from these problems that the standard distributive/collective classification is not tenable. He proposes to replace the traditional test in (3) with an alternative test that does not use definite plurals and conjunctions and works even when group nouns like *committee* are used in them. Winter's test leads to an alternative classification based on whether or not a predicate is sensitive to the distinction between singular quantificational determiners like *every* and plural ones like *all*. Distributive predicates like *smile* are compatible with both kinds of determiners and lead to equivalent interpretations. Winter calls this class *atom predicates* (see (7)). Some collective predicates, like *be numerous*, show the same behavior as distributive predicates like *smile*, while others like *gather*, which he calls *set predicates* (see (8)), distinguish between both.

- (7) **Atom predicates**
 - a. All the children smiled. \Leftrightarrow Every child smiled.
 - b. All of the enemy armies are numerous.
 \Leftrightarrow Every enemy army is numerous.¹¹

¹¹ This judgment is based on Kroch (1974: 194). For other speakers, *numerous* cannot be applied to singular group nouns. Krifka (2004) lists **The Jones family is numerous* as ungrammatical, with reference to Kleiber (1989). Readers who share this judgment may want to substitute other predicates, like *be large in number*, in order for the argument in the main text to go through.

TABLE 4.1. Comparison of the distributive-collective and atom-set typologies

Example	Traditional	Winter	This book
<i>smile</i>	distributive	atom predicate	distributive
<i>be numerous</i>	collective		<i>numerous</i> -type
<i>gather</i>		set predicate	<i>gather</i> -type

(8) **Set predicates**

- a. All the children gathered. \nRightarrow *Every child gathered.
- b. All the committees gathered. \nRightarrow Every committee gathered.

As shown in Table 4.1, Winter’s test draws the boundary at a different place than the traditional distributive-collective criteria. For this reason, it is not useful as a characterization of distributive predicates, which it is not meant to be. On the other hand, by placing the boundary within the traditional class of collective predicates, Winter’s test introduces a new and useful distinction within that class.

The categories I use are also shown in Table 4.1. They represent a synthesis of both the traditional categories and those of Winter. Distributive predicates are retained as a category, and collective predicates are split into *numerous*-type and *gather*-type predicates. Section 10.4 proposes to account for the differing behavior of these two classes of predicates in terms of distributivity to subgroups. For now, I leave collective predicates aside and concentrate on the notion of a distributive predicate.

Winter does not make any distinction between distributive and collective predicates because he does not consider this distinction well-motivated. At least for the purpose of this book, however, it is useful to have an operational definition of a distributive predicate. To develop such a definition, it is necessary to address Winter’s concerns about the reliability of the traditional tests. I propose to do that by slightly reformulating the tests in order to control for effects related to nonmaximal interpretations, conventionalized coordinations, and group nouns.

Indefinite numerals, particularly those involving small numbers, are not as likely as definite plurals to give rise to nonmaximal interpretations. For example, in a scenario like the one in (5), the entailment in (9a) can fail because of nonmaximality, but not the entailment in (9b).

- (9) a. The reporters spoke up. \nRightarrow Each of the reporters spoke up.
- b. Three reporters spoke up. \Leftrightarrow Three reporters each spoke up.

We can now refine the traditional definition of a distributive predicate in a way that avoids constructions involving coordinations and definite plurals:

(10) **Operational definition: Distributive predicate**

A distributive predicate is a predicate for which (11a) and (11b) are acceptable and entail each other when it is substituted for PRED.

- (11) a. Three people PRED.
b. Three people each PRED.

The noun *people* may be replaced by another noun if necessary to avoid selectional restrictions such as animacy requirements. To address Winter's concern about group nouns, we restrict the test by agreeing that this noun may not be replaced by a group noun. I propose an operational definition of group nouns in Section 2.6.4, which is based on a test proposed by Barker (1992).

To mention a few examples, the predicates *sleep*, *run*, *sneeze*, *get up*, *wear a dress*, and *take a breath* are all distributive predicates according to the definition in (10) because the entailment from (11a) to (11b) is obligatory with them. The predicates *eat a pizza*, *build a raft*, and *ask a question* are not distributive because the entailment from (11a) to (11b) is not obligatory with them. The predicates *meet* and *be numerous* are not distributive because at least (11b) is not acceptable, except (for some speakers) when we replace the word *people* by a group noun such as *committee*, which by convention is not allowed.

Following standard usage in the literature, I have described predicative distributivity as a property of intransitive predicates. The notion can be generalized to describe transitive predicates, but in this case it needs to be relativized to an argument position or thematic role. I come back to this idea in Section 4.5.

4.3 Distributive constructions

Section 4.2 has shown how distributivity can be seen as a property of predicates, of quantifiers, and of pairs of constituents. We can see distributivity in yet another way, namely as a property of entire constructions. By understanding distributivity in this way, we can abstract away from individual sentences and predicates. This allows us to generalize the concept of distributivity to encompass restrictions on aspectual properties and measure functions.

For this purpose, I introduce the following terminology. A *distributive construction* is a lexicosyntactic configuration that imposes an obligatory distributive relation between two of its constituents. A sentence *instantiates* a distributive construction when the sentence exemplifies the syntactic configuration of that construction. We have already seen that the adverbial-*each* construction obligatorily contains a distributive relation. Sentences with *each* are therefore uncontroversial examples of distributive constructions.

This section provides initial motivation for treating adverbial-*each* constructions, *for*-adverbials, and pseudopartitives as members of a natural class, namely the class of distributive constructions.

4.3.1 *For-adverbials are distributive constructions*

For-adverbials can be classified as distributive constructions because they obligatorily involve a distributive relation. On the assumptions in Section 3.3, a sentence like (12a) involves reference to an event e_0 whose runtime is an interval of five-minute length. This sentence entails (12b) and (12c).

- (12) a. John ran for five minutes.
 b. \Rightarrow John ran for four minutes.
 c. \Rightarrow John ran for three minutes.

In the present event-semantic framework, the entailments in (12) involve the existence of an event e_1 in which John ran for four minutes, an event e_2 in which he ran for three minutes, and so on. The sentence features a distributive relation because two of its constituents—the *for*-adverbial and the verbal or sentential predicate it modifies—jointly determine this entailment and because on the semantic assumptions in Chapter 2, the events e_1, e_2, \dots that represent these entailments are parts of the event e_0 .

4.3.2 *Pseudopartitives are distributive constructions*

Pseudopartitives, such as the subject of (13), can also be classified as distributive constructions.

- (13) Three liters of water are sufficient.

Because the word *water* occurs in (13), the sentence gives rise to the entailment that the parts of the quantity in question consist themselves of water. This entailment only involves one constituent, namely *water*. However, the sentence also has another distributive entailment: among these parts there exist one liter of water, two liters of water, and so on. This existence entailment is present despite the fact that the sentence has a nondistributive verb phrase, *be sufficient*, so the entailment must be due to the pseudopartitive rather than the verb phrase. It is a distributive entailment because it concerns the parts of an entity, namely the water entity to which the pseudopartitive refers, and it involves a distributive relation because the meanings of both the constituent *water* and the constituent *three liters* contribute to it. It is easy to see that the presence of this distributive relation is not accidental, but that it is required by the construction, because every pseudopartitive gives rise to similar entailments. Since pseudopartitives obligatorily involve a distributive relation, I classify them as distributive constructions.

4.4 The components of a distributive relation

Section 4.3 established that adverbial-*each* constructions, *for*-adverbials, and pseudopartitives are distributive constructions because they involve obligatory distributive

relations between their constituents. Here, I introduce terminology with which we can express generalizations over the three constructions.

As Zimmermann (2002b: 23) observes, there is considerable terminological confusion in the literature concerning the components of a distributive relation. Table 4.2, which is expanded from that source, gives an overview of the terminology.

For conciseness, I adopt the terms *Key* and *Share*, as in Gil (1989) and Choe (1991).¹² Intuitively, the term *Share* refers to the constituent whose denotation is distributed over the parts of the referent of the other constituent, which is called the *Key*. For example, the property of reading a book is distributed over the individual boys, the property of being water is distributed over the liters, and the property of being an event of pushing a cart is distributed over the hours. My use of the terminology is illustrated in Table 4.3, repeated here in modified form from Table 1.1 in Chapter 1. The constituent structure presupposed by this use is justified in Chapter 3.

The intuitive criterion behind the assignment of the terms *Key* and *Share* just described is perhaps not completely clear. I also use a formal criterion, which is more dependent on my specific choice of background assumptions, but which coincides with the intuitive criterion and makes this assignment clearer. As discussed in

TABLE 4.2. Terms for the components of a distributive relation

Source	Name for “ <i>The boys</i> ”	Name for “(<i>took</i>) a <i>breath</i> ”
Link (1998b)	distributional domain	Distributive Share (DstrShr)
Choe (1987)	Sorting Key (SrtKy)	Distributed Share (DstrShr)
Safir & Stowell (1988)	Range NP	Distributing NP (DistNP)
Gil (1989), Choe (1991)	Key	Share
Zimmermann (2002b)	DistKey	DistShare
Blaheta (2003)	Range	Dist phrase
This book	Key	Share

TABLE 4.3. A bridge from distributivity to aspect and measurement

Construction	Example	Key	Share	Map
Adverbial <i>each</i>	Three boys each laughed	three boys	laugh	agent
<i>For</i> -adverbial	John ran for three hours	three hours	John run	runtime
Pseudopartitive	three liters of water	three liters	water	volume

¹² The term “Sorting Key” in Choe (1987) is taken from Kuno (1982), who uses it in describing answers to multiple wh-questions. There are small differences not only in terminology but also in its application. Choe, and Safir & Stowell, use their terms *Distributed Share* and *DistNP* differently than other authors: they apply it to the noun phrase *a breath*, as opposed to the entire verb phrase *took a breath*. Link uses the term *Distributional Key* only for Keys that also exhibit quantificational distributivity, like *every*.

Section 2.5, I adopt a number of standard assumptions to the effect that constituents are related by certain covert functions: thematic roles, trace functions, and measure functions. These functions coincide with distributive relations in a particular way: they always map entities associated with the Share to entities associated with the Key. For example, in *Three boys each laughed*, the function *agent* maps events associated with the Share to entities associated with the Key. In *John ran for three hours*, the function *runtime* maps events associated with the Share to intervals associated with the Key; and in *three liters of water*, the function *volume* maps substances associated with the Share to degrees associated with the Key. It can be useful to refer to these functions independently of the construction involved. As shown in Table 4.3, I use the term *Map* for this purpose.

4.5 The constraints on distributive constructions

Recall that *each*-constructions only combine with distributive predicates, *for*-adverbials only combine with atelic predicates, and pseudopartitives accept only mass nouns and plural count nouns as substance nouns. Put in the terms just introduced, these facts all concern Shares: *each*-constructions accept only distributive predicates as Shares; *for*-adverbials accept only atelic predicates as Shares; and pseudopartitives accept only mass nouns and plural count nouns as Shares.

After presenting the traditional explanations of these facts, which are all formulated separately, I update each of them in successive refinements to reflect my own assumptions. In each case, the updates result in practically the same concept. This convergence forms a basis on which we can formulate a single constraint that bridges the differences between the three constructions.

4.5.1 Capturing predicative distributivity

Let me start by proposing a higher-order property, stratified distributive reference, which formally captures the notion of a distributive predicate that was operationally defined in Section 4.2.3. All distributive predicates can then be modelled as having stratified distributive reference. I will focus here on one-word predicates like *smile* or *sneeze*, which are distributive by virtue of their lexical semantics. Multi-word predicates like *take a breath* may become distributive as a result of the application of a distributivity operator, whose discussion I delay until Chapter 8. Stratified distributive reference provides a useful generalization that abstracts away from these different ways in which a predicate can be distributive. It encapsulates what it means to be a distributive predicate. Isolating this property is a crucial step on the way towards unifying it with its siblings in the domains of aspect and measurement.

Let me introduce stratified distributive reference in several steps. Our starting point is a relatively simple idea: we require that the only kinds of events in a distributive predicate are those which consist only of subevents to which the predicate also

applies and whose agent is an atom. The following preliminary definition encapsulates this idea.

(14) **Stratified distributive reference (preliminary definition 1)**

$$\text{SDR}_{\text{agent}}(P) \stackrel{\text{def}}{=} \forall e \left[P(e) \rightarrow \forall e' \leq e \left(\begin{array}{c} P(e') \wedge \\ \text{Atom}(*\text{agent}(e')) \end{array} \right) \right]$$

(An event predicate P has stratified distributive reference with respect to the thematic role *agent* iff every event e to which P applies only has subevents to which P also applies and whose agents are atoms.)

This preliminary definition does not work correctly, however. Even lexical predicates like *sneeze*, which are distributive according to the operational definition, do not satisfy it. For example, if e_1 is a sneezing event whose agent is John and e_2 is a sneezing event whose agent is Mary, then event $e_1 \oplus e_2$ has itself as a subevent, but its agent, $j \oplus m$, is not atomic, contrary to the requirement of the definition. For the definition to be true, each subevent of $e_1 \oplus e_2$ must have an atomic entity as its agent. The definition interferes with the assumption of lexical cumulativity (Section 2.7.2). This assumption has the consequence that the sum of two sneezing events is again in the denotation of *sneeze*. These sum events cause the definition to break down, because by cumulativity of thematic roles (Section 2.5.1), their agents are generally not atoms. These sum events have subevents whose agents are not atoms, namely themselves, because the subevent relation \leq is modeled by the mereological parthood relation, and is therefore reflexive (Section 2.3.1).

A first attempt at fixing the definition would consist in replacing \leq (parthood) with $<$ (proper parthood) in order to avoid the problem caused by reflexivity. However, this does not work either: if the toy model is extended by a third sneezing event e_3 , lexical cumulativity causes the event $e_1 \oplus e_2 \oplus e_3$ to be a sneezing event, and the event $e_1 \oplus e_2$ is a proper part of this event. Since $e_1 \oplus e_2$ does not have an atomic agent, the definition again does not apply to *sneeze*.

Another attempt would consist in using the relation $<_{\text{Atom}}$ (proper atomic parthood). Since e_1 and e_2 each have atomic agents, this attempt causes *sneeze* to have stratified distributive reference as desired, assuming that e_1 and e_2 are absolute atoms. This assumption is plausible because sneezing events are punctual. But there is no guarantee that distributive predicates always apply exclusively to events that consist of atomic events. Section 2.4.3 has already pointed out that it is not necessary to assume that singular events are atomic. It is in fact implausible to make this assumption, because non-punctual events in the extension of atelic predicates are plausibly not atomic. Some of these atelic predicates, such as *run*, are also distributive, but they would not be predicted to be distributive under this attempt.

These two attempts at extending the concept of a distributive predicate to an event-based framework have failed because of problems in selecting the right subevents with atomic agents. These problems can be avoided in a definition that remains neutral on

the issue of how to identify these subevents. Schematically, instead of requiring for any event e that *every* subevent (or proper subevent, or proper atomic subevent) of e has an atomic agent, we require that there is *some* way of dividing every event e into subevents with atomic agents. In (15), this is expressed by the star operator. To understand how this definition works, recall that $e \in * \lambda e'. C(e')$ is true iff there is a way to divide e into one or more possibly overlapping parts that are in C . To use a term from Schwarzschild (1996), C must essentially act as a cover of e . See Section 5.4 for more discussion of this point.

(15) **Stratified distributive reference (preliminary definition 2)**

$$\text{SDR}_{\text{agent}}(P) \stackrel{\text{def}}{=} \forall e \left[P(e) \rightarrow e \in * \lambda e' \left(\begin{array}{c} P(e') \wedge \\ \text{Atom}(*\text{agent}(e')) \end{array} \right) \right]$$

(An event predicate P has stratified distributive reference with respect to a thematic role *agent* iff every event e to which P applies can be exhaustively divided into one or more subevents (“strata”) to which P also applies and whose agent is an atom.)

On this definition, *sneeze* is correctly predicted to have stratified distributive reference with respect to the agent role. This is so because e_1 , e_2 , and their sum $e_1 \oplus e_2$ can each be divided into one or more sneezing events with an atomic agent. The definition in (15) does not involve the notion of an atomic event, and this makes it possible for arguably nonatomic distributive predicates like *run* to have stratified distributive reference.

The definition is still preliminary, because it is tied to the thematic role *agent*. In the next step, we generalize it to arbitrary thematic roles θ . Definition (16) does so by replacing *agent* with θ .

(16) **Stratified distributive reference (universal version)**

$$\text{SDR}_{\theta}(P) \stackrel{\text{def}}{=} \forall e \left[P(e) \rightarrow e \in * \lambda e' \left(\begin{array}{c} P(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right) \right]$$

(An event predicate P has stratified distributive reference with respect to a thematic role θ iff every event e to which P applies can be exhaustively divided into one or more subevents (“strata”) to which P also applies and whose θ is an atom.)

This definition makes it easy to account for the inferential behavior of distributive predicates. I assume that our knowledge about whether or not a verb is distributive is part of what we know about its meaning. To mention a few examples, we know that whenever a plurality of individuals see or are seen, each of them sees or is seen. Thus, it follows both from *John and Mary saw Bill* and from *John saw Bill and Mary* that John saw Bill. This means that the verb *see* is lexically distributive on both its agent and theme positions. By contrast, the verb *lift* is not distributive on its agent position: from *John and Mary lifted box B* it does not follow that *John lifted box B*. The verb *kill*

is distributive on its theme but not on its agent (Lasersohn 1988, Landman 1996). This is illustrated in the following scenario. The two outlaws Bonnie and Clyde were killed by a posse of police officers, which included Sheriff Jordan. Given this background knowledge, (17a) entails (17b) but does not entail (17c), because Sheriff Jordan's actions might not have been sufficient by themselves to kill anyone.

- (17) a. The police officers killed the two outlaws.
 b. \Rightarrow Bonnie was killed.
 c. \nRightarrow Sheriff Jordan killed someone.

As a means of capturing the inferential properties of verbs, I will follow the approach of Hoeksema (1983: 68), who suggests that the difference between distributive and nondistributive predicates

is clearly a lexical matter, having to do with the way reality seems to be organized: two persons may buy a car together, without each of them buying a car, but if two persons are sleeping together, each must sleep individually. We may handle such issues by meaning postulates, which put restrictions on admissible models. Such meaning postulates may take the form of statements that predicate X is distributive, or not distributive . . .

I will formulate meaning postulates using stratified distributive reference. A typical meaning postulate will be relativized to a thematic role θ , and will state that whenever a (θ -)distributive predicate applies to (an event whose θ is) a plurality of individuals, it also applies to (events whose θ s are) all the individuals in the plurality. When a verb is distributive on more than one of its positions, we may formulate a different meaning postulate for each position. Take for example the verb *see*. Its lexical entry, shown in (18), includes a star operator as a reminder of the assumption that it is closed under sum (see Section 2.7.2):

- (18) **Lexical entry for *see***
 $\llbracket \text{see} \rrbracket = \lambda e. * \text{see}(e)$
 (The meaning of “see” is the property that holds of any sum of one or more seeing events.)

The meaning postulates in (19) and (20) capture the fact that the predicate denoted by the verb *see* is distributive on its theme and agent positions:

- (19) **Meaning postulate: *see* is distributive on its theme position**

$$\text{SDR}_{\text{theme}}(*\text{see}) \Leftrightarrow \forall e \left[\text{see}(e) \rightarrow e \in * \lambda e' \left(\begin{array}{c} * \text{see}(e') \wedge \\ \text{Atom}(\text{theme}(e')) \end{array} \right) \right]$$

- (20) **Meaning postulate: *see* is distributive on its agent position**

$$\text{SDR}_{\text{agent}}(*\text{see}) \Leftrightarrow \forall e \left[\text{see}(e) \rightarrow e \in * \lambda e' \left(\begin{array}{c} * \text{see}(e') \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right) \right]$$

The meaning postulate in (19) can be read as follows. Any seeing event e consists of one or more seeing events e' whose themes are atoms. For example, if e is an event in which John and Bill were seen, that event consists of (in this case at least two) seeing events whose themes are atoms. From this and from the assumption that John and Bill are mereological atoms, one can conclude that John was seen and that Bill was seen. In effect, (19) states that a sum of individuals can be seen only when each of them is seen. Likewise, (20) states that a sum of individuals can see something only when each of them sees something. We can capture the difference between the agent and theme role of *kill* by adopting a meaning postulate analogous to (19) only for the theme position of that verb, or by explicitly negating it for the agent position:

- (21) **Meaning postulates: *kill* is distributive with respect to themes but not agents**
 a. $\text{SDR}_{\text{theme}}(\text{kill})$
 b. $\neg \text{SDR}_{\text{agent}}(\text{kill})$

These postulates, together with the background assumptions on event semantics and lexical cumulativity, correctly account for the inferential behavior of *kill*: together with the translation of (22a) and appropriate formalizations of background knowledge such as the fact that Sheriff Jordan is an officer, as in (22b), they entail the translation of (22c), but not that of (22d). For ease of exposition, I represent the agents of (22a) and (22c) as sums rather than groups. The difference does not matter here since there are no inferences on the agents in these sentences.

- (22) a. The officers killed Bonnie and Clyde.
 $\exists e[*\text{kill}(e) \wedge *\text{agent}(e) = \bigoplus \text{officer} \wedge *\text{theme}(e) = b \oplus c]$
 b. Sheriff Jordan is an officer.
 $\text{officer}(j)$
 c. \Rightarrow Bonnie was killed.
 $\exists e'[\text{kill}(e') \wedge \text{theme}(e') = b]$
 d. \nRightarrow Sheriff Jordan killed someone.
 $\exists e''[\text{kill}(e'') \wedge \text{agent}(e'') = j]$

The approach I am adopting here allows and requires us to make a clear distinction between the lexical entry of a word and what we know about the events it describes. I do not take (19) and (20) to be part of the lexical entry for *see*. This division of labor between lexical entries like (18) and higher-order properties like the ones encoded in (19) and (20) is a hallmark of algebraic semantics.

I have called the definition of stratified distributive reference in (16) “universal” because it quantifies over all events in the denotation of the predicate. This is appropriate as a way to formulate meaning postulates in some cases. But in other cases, it is too stringent. Consider predicates like *meet*, which we have classified as collective because they are normally incompatible with *each*, as shown in (23a).

- (23) a. *Three boys each met.
 b. Three committees each met.

As (23b) shows, group nouns such as *committee* are an exception to this generalization. In Section 4.2.3, this led us to exclude them from the operational definition of distributive predicates. Definition (16) cannot distinguish between (23a) and (23b) because it is not sensitive to the nature of the predicate involved. A plurality of boys can be the (nonatomic) agent of a meeting event which does not consist of meeting subevents with atomic agents. Therefore *meet* does not have stratified distributive reference with respect to agents. What does not enter the picture on definition (16) is the difference between boys and committees. By lexical cumulativity, whenever there are three committees and each of them is the agent of a meeting event, their sum is also the agent of a meeting event. This plurality of committees is the agent of a meeting event which consists of meeting subevents with atomic agents, namely the individual committees. (As described in Section 2.6.4, I assume with Barker (1992) that the entities in the denotation of the word *committee* are atoms.) Intuitively, we may say that *meet* is distributive with respect to certain events (like sums of meeting events whose agents are committees) but not with respect to others (like meeting events whose agents are sums of boys). In order to express this, we need to restrict stratified distributive reference to individual events. This can be done by supplying the event in question to stratified distributive reference as an additional argument:¹³

(24) **Stratified distributive reference (restricted version)**

$$\text{SDR}_\theta(P)(e) \stackrel{\text{def}}{=} e \in * \lambda e' \left(\begin{array}{c} P(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$$

(An event predicate P has stratified distributive reference with respect to a thematic role θ and an event e iff e can be exhaustively divided into one or more subevents (“strata”) to which P also applies and whose θ is an atom.)

I write $\text{SDR}_\theta(P)(e)$ rather than $\text{SDR}_\theta(P, e)$ to hint at the fact that one may λ -abstract over one or both arguments of this property, as in $\text{SDR}_\theta(P) = \lambda e. \dots$ or $\text{SDR}_\theta = \lambda P \lambda e. \dots$. This means we may think of restricted SDR and related notions as higher-order properties or as modifiers of predicates. In Chapter 8, I will exploit this fact to propose a reformulation of various distributivity operators in terms of restricted SDR and related notions.

Since meaning postulates formalize aspects of our world knowledge that interacts with semantics, it is difficult to give clear criteria as to how specifically or how generally they should be formulated. Instead of futilely attempting to systematize our world knowledge, I will limit myself to illustrating the general strategy by sketching a meaning postulate that is specific to meetings and committees. One could easily

¹³ I am grateful to Piñón (2015) and Schwarzschild (2015) for suggesting this restriction in another context.

broaden its scope by formally specifying appropriate classes of collective predicates and group nouns.

- (25) **Meaning postulate: *meet* is distributive on its theme position for events whose agents are sums of committees**

$$\forall e. *meet(e) \wedge *committee(*agent(e)) \rightarrow SDR_{theme}(*meet)(e)$$

Based on the definition in (24), this meaning postulate expands as follows:

- (26) $\forall e. *meet(e) \wedge *committee(*agent(e)) \rightarrow$

$$e \in *_{\lambda e'} \left(\begin{array}{c} *meet(e') \wedge \\ Atom(theme(e')) \end{array} \right)$$

(Any sum of meeting events whose plural agent is a sum of committees consists of meeting events whose agents are individual committees.)

4.5.2 *Adverbial each*

Now that we have the means to formally specify what it means for a predicate to be distributive, we can formally specify what it means for a construction to require its Share to be distributive. Let me start with the adverbial-*each* construction. By the operational definition in (10), any distributive predicate can be the Share of this construction. As we have seen, collective predicates can also be used in that capacity as long as they can be interpreted distributively, which requires group nouns like *committee*. With the help of restricted SDR, we can now express this fact formally:

- (27) **Constraint on Shares of adverbial-*each* sentences**

A sentence with adverbial *each* whose Share is *S*, whose Map is *M*, and which is used to describe an event *e* is acceptable only if *S* has stratified distributive reference with respect to *M* and *e* (formally: $SDR_M(S)(e)$).

The following examples illustrate how this constraint works. In each case, I assume that the event in question is quantified over by an existential closure operator (see Section 2.10) and I will refer to it as *e*.

- (28) Three boys each laughed.

- a. Key: Three boys
- b. Map: agent
- c. Share: laugh

Applied to example (28), constraint (27) predicts that the sentence is acceptable only if the condition $SDR_{agent}(*laugh)(e)$ is fulfilled, where *e* is the event that verifies the sentence. This condition expands as follows:

(29) $\text{SDR}_{\text{agent}}(\text{laugh})(e)$

$$\Leftrightarrow e \in * \lambda e' \left(\begin{array}{c} \text{laugh}(e') \wedge \\ \text{Atom}(*\text{agent}(e')) \end{array} \right)$$

(The event e can be divided into one or more laughing events (“strata”) whose agent is an atom.)

This condition is fulfilled on the assumption that stratified distributive reference accurately models predicative distributivity, i.e. on the assumption that distributive predicates (according to the operational definition) have stratified distributive reference (according to the formal definition). This assumption is independently needed to explain the inference from (30a) to (30b).

- (30) a. John and Mary laughed.
b. \Rightarrow John laughed.

I formalize the assumption that *laugh* is distributive on its agent position by formulating a meaning postulate that is analogous to the one in (20):

(31) **Meaning postulate: *laugh* is distributive on its agent position**

$$\text{SDR}_{\text{agent}}(*\text{laugh}) \Leftrightarrow$$

$$\forall e \left[\text{laugh}(e) \rightarrow e \in * \lambda e' \left(\begin{array}{c} *\text{laugh}(e') \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right) \right]$$

This meaning postulate is formulated in terms of universal SDR, which is a stronger notion than restricted SDR:

(32) **Fact: Universal SDR entails restricted SDR**

$$\forall P \forall \theta. \text{SDR}_{\theta}(P) \rightarrow \forall e. [P(e) \rightarrow \text{SDR}_{\theta}(P)(e)]$$

(If a predicate has universal SDR, then for any event in its denotation, it has restricted SDR relative to that event.)

Together with the meaning postulate in (31), this entails that $\text{SDR}_{\text{agent}}(*\text{laugh})(e)$ is fulfilled for any event e that verifies the sentence.

Let us now consider the case of a nondistributive predicate:

- (33) *Three boys each met.
a. Key: Three boys
b. Map: agent
c. Share: meet

Applied to example (33), constraint (27) predicts that the sentence is acceptable only if the condition $\text{SDR}_{\text{agent}}(\text{meet})(e)$ is fulfilled, where e is the event that verifies the sentence. This condition expands as follows:

(34) $\text{SDR}_{\text{agent}}(\text{meet})(e)$

$$\Leftrightarrow e \in * \lambda e' \left(\begin{array}{c} \text{meet}(e') \wedge \\ \text{Atom}(*\text{agent}(e')) \end{array} \right)$$

(The event e can be divided into one or more meeting events whose agent is an atom.)

Any event e that would verify (33) would also have a plural individual corresponding to three boys as its agent. Such an event would not be able to satisfy the condition in (34). A plural individual corresponding to three boys can meet, and this meeting event will not have any parts whose agent is an atom. This assumption is independently needed to explain the lack of inference from (35a) to (35b).

- (35) a. John and Mary met.
b. \nRightarrow *John met.

Finally, the following example illustrates what happens when *meet* is applied to a group noun:

- (36) Three committees each met.
a. Key: Three committees
b. Map: agent
c. Share: meet

Applied to example (36), constraint (27) predicts that the sentence is acceptable only if the condition $\text{SDR}_{\text{agent}}(\text{meet})(e)$ is fulfilled, where e is the event that verifies the sentence. This condition expands to (34) as in the previous example, but now e is a different kind of event than before. Any event that verifies (36) will have a plural individual corresponding to three committees as its agent. Unlike in the previous case, such an event can satisfy the condition in (34) if there are three committees and each of them met. Now e can be the sum of the three meeting events. By the meaning postulate in (25), each of the three events is a meeting event with a committee as its agent, and these committees are atoms. Therefore e can verify (36).

4.5.3 Temporal for-adverbials

The Share of a temporal *for*-adverbial must be an atelic predicate. As Dowty (1979) shows, the entailment properties of atelic predicates with respect to a large range of other phenomena, including tense, the progressive, and aspectualizers, can be represented as a first approximation by modeling the denotations in terms of the subinterval property. This is an old idea. For example, Vendler (1957) suggests that activity predicates like *run* have the property that “any part of the process is of the same nature as a whole.” He contrasts this property with accomplishment predicates like *run a mile*, which “proceed toward a terminus which is logically necessary to their being what they are.” The model-theoretic implementation of the subinterval property

is usually attributed to Bennett & Partee (1972). In their original formulation, matrix verb phrases with this property are true at every subinterval, including every moment, of the interval at which the sentence that contains them is true. In our event-based framework, the subinterval property can be defined with the help of the relation AT from Section 2.5.2:

(37) **Definition: Subinterval property**

$$\text{SUB}(P) \stackrel{\text{def}}{=} \forall i[\text{AT}(P, i) \rightarrow \forall j[j < i \rightarrow \text{AT}(P, j)]]$$

$$\Leftrightarrow \forall i \forall e[P(e) \wedge \tau(e) = i \rightarrow \forall j[j < i \rightarrow \exists e'[P(e') \wedge \tau(e') = j]]]$$

(An event predicate P has the subinterval property iff whenever it holds at an interval, it also holds at every one of its subintervals.)

It is well known that the subinterval property is too strong. This is known as the minimal-parts problem (to be discussed in more detail in Chapter 5). For example, *waltz* is an atelic predicate, but a subinterval that is so short that it contains less than three steps makes it hard to determine whether these steps qualify as waltzing (Dowty 1979). For this reason, it is difficult to know whether the subinterval property in (37) applies to *waltz*, and this makes it unsuited for modeling atelicity.

Definition (38) implements a weaker version of the subinterval property that avoids the minimal-parts problem:

(38) **Definition: Stratified subinterval reference**

$$\text{SSR}(P)(e) \stackrel{\text{def}}{=} e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \tau(e') < \tau(e) \end{array} \right)$$

(An event predicate P has stratified subinterval reference with respect to an event e iff e can be exhaustively divided into parts (“strata”) that are each in P and whose runtimes are properly included in the runtime of e .)

The exact formulation of this definition is motivated in Chapter 5, but the basic intuition behind it is the following:

- Instead of testing whether the predicate holds at every subinterval of the runtime of the event in question, the definition only tests whether that runtime can be divided into smaller subintervals.¹⁴
- Instead of quantifying over all events to which the predicate applies, the definition only quantifies over all the subevents of a given event e . Intuitively, this is the event that the sentence in question is about.¹⁵

¹⁴ More precisely, the definition tests whether the event can be divided into subevents whose runtimes are proper subintervals of its own runtime. This difference is not crucial. Since the runtime function τ is a sum homomorphism (Section 2.5.2), the interval $\tau(e)$ is guaranteed to be the sum of all the runtimes of the parts of e .

¹⁵ This restriction was suggested by Piñón (2015) and Schwarzschild (2015) in order to avoid making claims about unrelated events.

- By testing whether the event can be subdivided into events with shorter runtimes, I avoid explicit reference to atomic events or intervals, similarly to the definition of stratified distributive reference in (16).¹⁶

We can now formulate the constraint on Shares in temporal *for*-adverbials with the help of stratified subinterval reference.

(39) **Constraint on Shares of temporal *for*-adverbials**

A temporal *for*-adverbial whose Share is S and which is used to describe an event e is acceptable only if S has stratified subinterval reference with respect to e (formally: $\text{SSR}(S)(e)$).

The examples in (40) illustrate how this constraint works:

- (40) John and Mary waltzed for an hour.
- a. Key: an hour
 - b. Map: runtime
 - c. Share: waltz

Applied to example (40), constraint (39) predicts that the sentence is acceptable only if the condition $\text{SSR}(\text{waltz})(e)$ is fulfilled with respect to some event e in the denotation of *waltz*. This condition expands as follows:

- (41) $\text{SSR}(\text{waltz})(e)$
 $\Leftrightarrow e \in {}^*\lambda e' \left(\begin{array}{l} \text{waltz}(e') \wedge \\ \tau(e') < \tau(e) \end{array} \right)$
 (The event e can be divided into waltzing events whose runtime is properly included in its own.)

This condition is fulfilled by any sufficiently long event in the denotation of the predicate *waltz*, as shown by the inference patterns like (42) (Dowty 1979).

- (42) a. John and Mary waltzed from 3pm to 4pm.
 b. \Rightarrow John and Mary waltzed from 3pm to 3:30pm.

The example in (43) illustrates the case of the telic predicate *collapse*:

¹⁶ I adopted a more stringent condition in Champollion (2010b, 2015c), where instead of merely requiring the subevents to have shorter runtimes than that of e , I required their runtimes to be very small in comparison to the runtime of e . What this meant was specified through a threshold $\varepsilon(K)$, which I did not formally define. This threshold was linked to the Key (the temporal noun phrase) of the *for*-adverbial. I assumed that ε was a predicate which takes a comparison class K and an entity x and returns true iff x is very small with respect to the comparison class K . Piñón (2015) and Schwarzschild (2015) argue convincingly that this is unnecessary, even though in practice it may be the case that many predicates that combine with *for*-adverbials do satisfy the more stringent condition. Metaphorically speaking, if *for*-adverbials are a sieve, the holes of the sieve can be fairly large even though most of the grains of sand that pass through it tend to be small. For more discussion of this modification of the original theory, see Champollion (2015b).

- (43) *The building collapsed for ten minutes.
- a. Key: ten minutes
 - b. Map: runtime
 - c. Share: collapse

Applied to example (43), constraint (39) predicts that the sentence is acceptable only if the condition $\text{SSR}(\text{collapse})(e)$ is fulfilled with respect to some event e in the denotation of *collapse*. This condition expands as follows:

- (44) $\text{SSR}(\text{collapse})(e)$
- $$\Leftrightarrow \forall e \left[\text{collapse}(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{c} \text{collapse}(e') \wedge \\ \tau(e') < \tau(e) \end{array} \right) \right]$$
- (The event e can be divided into one or more collapsing events whose runtime is properly included in its own.)

This condition is not fulfilled by any event in the denotation of *collapse* (Krifka 1998). More generally, no collapsing event can be divided into collapsing events whose runtimes are properly included in its own.

4.5.4 Pseudopartitives

The Share (the substance noun) of a pseudopartitive must be a mass term or a plural count term. As is often noted (e.g. Bach 1986), the subinterval property, which has been used to characterize atelicity, is similar to the concept of divisive reference, which has been used to characterize mass terms and plurals.¹⁷ The definition of divisive reference is repeated here from Section 2.3.5:

- (45) **Definition: Divisive reference**
- $$\text{DIV}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$$
- (A predicate P over ordinary objects has divisive reference iff every part of any object in its denotation is also in its denotation.)

As a characterization of Shares in pseudopartitives, there is a serious problem with divisive reference: all mass terms are acceptable Shares in pseudopartitives, but the near-consensus view in the semantic literature is that they do not all have divisive reference. The evidence for this view is reviewed in Section 5.2.2. The question then arises what kind of property can characterize the Shares of pseudopartitives.

An example of a mass noun that does not have divisive reference is the mass noun *cable*.¹⁸ It does not have divisive reference because not every part of a cable segment

¹⁷ I follow Krifka (1989a) in calling this concept *divisive reference*. Some authors refer to this property as *distributive reference*, but I have already used a similar term in (16) with a different meaning.

¹⁸ I refer to the mass noun *cable*, in the sense in which it appears in *a spool of cable* or *a lot of cable*. I refer to entities in the denotation of the mass noun *cable* as *cable segments*. The count noun use of *cable* is irrelevant here because singular count nouns do not occur as Shares of pseudopartitives. As described

qualifies as a cable segment itself. Linear segments of a cable segment do, but slices which run the length of the spool or which are taken through the middle along the length of the cable segment do not.

Schwarzschild (2002, 2006) observes that these facts correlate with the interpretation of pseudopartitives whose substance noun is *cable*: (46) can only be interpreted as being about a cable segment whose length is three inches, not as a cable segment whose diameter is three inches.

(46) three inches of cable

As discussed in Sections 3.2 and 4.4, I adopt the standard assumption that the measure and substance noun of a pseudopartitive are related by a covert measure function, which I call the Map. The different potential interpretations of (46) are assigned different logical representations which differ only in the choice of this measure function:

(47) $\lambda x[\text{cable}(x) \wedge \exists d[\text{length}(x) = d \wedge \text{inches}(d) = 3]]$
(available)

(48) $\lambda x[\text{cable}(x) \wedge \exists d[\text{diameter}(x) = d \wedge \text{inches}(d) = 3]]$
(unavailable)

These logical representations give us an initial handle on the cable problem. The task consists of finding a constraint that rules out (48) but not (47). Note that the use of the measure function *diameter* is not itself the problem, as there are other pseudopartitives which allow for an interpretation in which the covert measure function is *diameter*. Schwarzschild (2002) considers the scenario of a growing pool of oil seeping out of the ground. He observes that we can report its progress by declaring there to be first ten inches of oil, then fifteen inches of oil, and so on. As long as the measure function *diameter* is clearly retrievable from the context of the utterance, these pseudopartitives are easily understood as involving reference to the diameter of the pool.

I now propose a constraint that rules out (48), and I formulate it in such a way that it also excludes mass nouns and plural count nouns. Consider a cable segment whose length is ten feet and whose diameter is ten inches. Among its parts, there are segments whose length is three feet, others whose length is six feet, and so on. These segments each qualify as cable, and their length is always lower than the length of the entire segment. Their diameter, however, is always the same as the diameter of the entire segment.

in Section 2.6, I assume that there are two lexical entries for the count noun and the mass noun use of words like *cable*, and that singular count nouns have quantized reference. As described in Section 2.6.1, a contextual parameter specifies the set of cable segments which are in the denotation of the count noun *cable*.

These facts can be modeled either by considering only certain subsets of the parthood relation among the parts of the cable segment, or by loosening the requirement of divisive reference so as to take a dimension parameter into account. The former route is taken by Schwarzschild (2002, 2006), but it involves making the parthood relation dependent on context. I argue against this strategy in Chapter 7. For this reason, I take the latter route, which relies only on the mereological parthood relation. The idea is to relativize the concept of divisive reference with respect to a certain dimension or measure function such as length or diameter. Intuitively, the mass noun *cable* is divisive along its length, but not along its diameter. The concept in (49) realizes this intuition by adding a parameter μ for measure functions to the concept of divisive reference.

(49) **Stratified measurement reference (preliminary definition)**

$$\text{SMR}_\mu(P) \stackrel{\text{def}}{=} \forall x \left[P(x) \rightarrow \forall d \left[d < \mu(x) \rightarrow \exists y \left(\begin{array}{c} y < x \wedge P(y) \\ \wedge \mu(y) = d \end{array} \right) \right] \right]$$

(A predicate P has stratified measurement reference relative to a measure function μ iff for any object x to which P applies, every value smaller than the μ -value of x is the μ -value of some part of x to which P again applies.)

The equations in (50) illustrate the application of this definition to the cable example.

(50) $\text{SMR}_{\text{length}}(\text{cable}) \Leftrightarrow$

$$\forall x \left[\text{cable}(x) \rightarrow \forall d \left[d < \text{length}(x) \rightarrow \exists y \left(\begin{array}{c} y < x \wedge \text{cable}(y) \\ \wedge \text{length}(y) = d \end{array} \right) \right] \right]$$

(Every cable segment of length 10in. has as part a cable segment of length 9in., another one of length 8in., etc.)

The fact that cable segments do not contain cable segments with a smaller diameter can also be expressed:

(51) $\neg \text{SMR}_{\text{diameter}}(\text{cable}) \Leftrightarrow$

$$\neg \forall x \left[\text{cable}(x) \rightarrow \forall d \left[d < \text{diameter}(x) \rightarrow \exists y \left(\begin{array}{c} y < x \wedge \text{cable}(y) \\ \wedge \text{diameter}(y) = d \end{array} \right) \right] \right]$$

(Not every (in fact, no) cable segment of diameter 10in. has as part a cable segment of diameter 9in., etc.)

While the preliminary definition of SMR in (49) successfully links the reference properties of *cable* with the available measure functions in pseudopartitives, it does not yet address the fundamental problem of divisive reference. As discussed in Chapter 5, the substances in the denotation of many mass nouns have parts outside of their denotations. This problem is analogous to the minimal-parts problem of atelic predicates that we encountered in the previous section. For example, *cable* is not strictly speaking divisive, because very short slices taken from a cable segment may not

qualify as cable segments. The solution to this problem is also analogous to the case of *for*-adverbials: instead of testing that *all* values below the μ -value of the substance are themselves μ -values of parts of the entity, we merely require that the substance in question can be divided into parts with smaller μ -values. The concept of stratified measurement reference is therefore redefined in (53). This definition tests whether there is a way of dividing the substance x into parts in the denotation of a predicate P which have a smaller μ -value.

(52) **Final definition: Stratified measurement reference (universal version)**

$$\text{SMR}_{\mu}(P) \stackrel{\text{def}}{=} \forall x \left[P(x) \rightarrow x \in {}^*\lambda y \left(\begin{array}{c} P(y) \wedge \\ \mu(y) < \mu(x) \end{array} \right) \right]$$

(A predicate P over entities has stratified measurement reference with respect to a function μ iff every entity x to which it applies can be exhaustively divided into parts (“strata”) that are each in P and whose μ -values are smaller than the μ -value of x .)

Finally, as in the case of SDR and SSR, we can relativize SMR to a specific entity x :

(53) **Final definition: Stratified measurement reference (restricted version)**

$$\text{SMR}_{\mu}(P)(x) \stackrel{\text{def}}{=} x \in {}^*\lambda y \left(\begin{array}{c} P(y) \wedge \\ \mu(y) < \mu(x) \end{array} \right)$$

(A predicate P over entities has stratified measurement reference with respect to a function μ and a substance x iff x can be exhaustively divided into parts (“strata”) that are each in P and whose μ -values are smaller than the μ -value of x .)

For example, the statement that the predicate *cable* has SMR with respect to length and a given cable segment x amounts to the statement that this cable segment can be divided into parts of shorter length that each qualify as cable.

We can now formulate the constraint on Shares in pseudopartitives with the help of the restricted version of stratified measurement reference.

(54) **Constraint on Shares of pseudopartitives**

A pseudopartitive whose Share is S , whose Map is M , and which is used to describe an entity x is acceptable only if S has stratified measurement reference with respect to M and x (formally: $\text{SMR}_M(S)(x)$).

The examples in (55) illustrate how this constraint works. The first example assumes that the covert Map is length:

- (55) three inches of cable
 a. Key: three inches
 b. Map: length
 c. Share: cable

Applied to example (55), constraint (54) predicts that a sentence in which (55) is used to describe an entity x is acceptable only if the condition $\text{SMR}_{\text{length}}(\text{cable})(x)$ is fulfilled. This condition expands as follows:

$$(56) \quad \text{SMR}_{\text{length}}(\text{cable})(x) \Leftrightarrow \\ x \in {}^*\lambda y \left(\begin{array}{l} \text{cable}(y) \wedge \\ \text{length}(y) < \text{length}(x) \end{array} \right) \\ (\text{x can be divided into cable segments ("strata") which are shorter than } x \text{ is.})$$

This condition is fulfilled in the case of (55). We started by assuming that *cable* has divisive reference relative to length. We then concluded that this assumption is too strong because arbitrarily short parts of a cable segment do not always qualify as cable segments. The new definition corrects this problem because it merely requires that whichever cable segment is described by (55) can be divided into shorter cable segments. That cable segment is three inches long, so there will be no problem to find shorter cable segments that are parts of it.

The Example (57) illustrates why *cable* cannot be used in a pseudopartitive in connection with a diameter Map:

- (57) three inches of cable
 a. Key: three inches
 b. Share: cable
 c. Map: diameter

Applied to example (57), constraint (54) predicts that the sentence is acceptable only if the condition $\text{SMR}_{\text{diameter}}(\text{cable})(x)$ is fulfilled. This condition expands as follows:

$$(58) \quad \text{SMR}_{\text{diameter}}(\text{cable})(x) \Leftrightarrow \\ x \in {}^*\lambda y \left(\begin{array}{l} \text{cable}(y) \wedge \\ \text{diameter}(y) < \text{diameter}(x) \end{array} \right) \\ (\text{x can be divided into cable segments ("strata") which have a smaller diameter than } x \text{ does.})$$

This condition is not fulfilled, because when a cable segment is cut lengthwise, the resulting parts do not generally qualify as cable segments.

4.6 Unifying the constraints

The fundamental parallel across distributive constructions emerges when we put the properties that were defined in Section 4.5 side by side:

(59) **Definition: Stratified distributive reference**

$$\text{SDR}_{\theta}(P)(e) \stackrel{\text{def}}{=} e \in {}^*\lambda e' \left(\begin{array}{l} P(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right) \quad = (24)$$

(60) **Definition: Stratified subinterval reference**

$$\text{SSR}(P)(e) \stackrel{\text{def}}{=} e \in {}^*\lambda e' \left(\begin{array}{l} P(e') \wedge \\ \tau(e') < \tau(e) \end{array} \right) = (38)$$

(61) **Definition: Stratified measurement reference**

$$\text{SMR}_\mu(P)(x) \stackrel{\text{def}}{=} x \in {}^*\lambda y \left(\begin{array}{l} P(y) \wedge \\ \mu(y) < \mu(x) \end{array} \right) = (53)$$

We arrived at these three properties from very different starting points: stratified distributive reference through the attempt to model predicative distributivity in an event-based framework; stratified subinterval reference through a refinement of the subinterval property that takes the minimal-parts problem into account; and stratified measurement reference through the attempt to explain why *three inches of cable* cannot refer to a three-inch-thick cable segment.

Although the three properties were independently motivated, they are very similar to each other. They can now be subsumed under one common property, which I simply call *stratified reference*.

(62) **Definition: Stratified reference (restricted version)**

Let d (a “dimension”) be any function from entities of type α to entities of type β , and let g (a “granularity level”) be any predicate of entities of type β . Let P range over predicates of type $\langle \alpha, t \rangle$ where α is either e or v , and let x range over entities of type α . Then:

$$\text{SR}_{d,g}(P)(x) \stackrel{\text{def}}{=} x \in {}^*\lambda y \left(\begin{array}{l} P(y) \wedge \\ g(d(y)) \end{array} \right)$$

(A predicate P (the “Share”) stratifies x with respect to a function d (the “dimension” or “Map”) and a predicate g (the “granularity level”) iff x can be exhaustively divided into parts (“strata”) which are each in P and which are each mapped by d to something in g .)

By quantifying over all values of x , we obtain the universal version of stratified reference:

(63) **Definition: Universal stratified reference**

$$\text{SR}_{d,g}(P) \stackrel{\text{def}}{=} \forall x [P(x) \rightarrow \text{SR}_{d,g}(P)(x)]$$

(P has universal stratified reference with dimension d and granularity g iff P stratifies everything in P .)

In expended form, this definition looks as follows:

(64) **Definition: Universal stratified reference (expanded)**

$$\text{SR}_{d,g}(P) \stackrel{\text{def}}{=} \forall x \left[P(x) \rightarrow x \in {}^*\lambda y \left(\begin{array}{l} P(y) \wedge \\ g(d(y)) \end{array} \right) \right]$$

(P has universal stratified reference along dimension d with granularity g iff any x in P can be divided into one or more parts in P that are each mapped by d to something in g .)

I used universal stratified reference in earlier work (Champollion 2010b, 2015c) and I will use it in Chapter 10 for reasons explained there. Following arguments by Piñón (2015) and Schwarzschild (2015), I will use the restricted version in (62) in the rest of this book. For more discussion of this point, see Champollion (2015b).

The symbols for constants and variables in these definitions should be taken to be unsorted, since the definition is intended to subsume properties of event predicates and properties of predicates over ordinary objects.

With the help of stratified reference, we can now unify the constraints on distributive constructions which were formulated separately in (27), (39), and (54), and which are repeated here for convenience.

(65) **Constraint on Shares of adverbial-*each* sentences** = (27)

A sentence with adverbial *each* whose Share is S , whose Map is M , and which is used to describe an event e is acceptable only if S has stratified distributive reference with respect to M and e (formally: $\text{SDR}_M(S)(e)$).

(66) **Constraint on Shares of temporal *for*-adverbials** = (39)

A temporal *for*-adverbial whose Share is S and which is used to describe an event e is acceptable only if S has stratified subinterval reference with respect to e (formally: $\text{SSR}(S)(e)$).

(67) **Constraint on Shares of pseudopartitives** = (54)

A pseudopartitive whose Share is S , whose Map is M , and which is used to describe an entity x is acceptable only if S has stratified measurement reference with respect to M and x (formally: $\text{SMR}_M(S)(x)$).

With the understanding that the adverbial-*each* constructions, pseudopartitives, and *for*-adverbials are distributive constructions, these constraints can now be replaced by the statement in (68):

(68) **Distributivity Constraint**

A distributive construction whose Share is S , whose Map is M , and which is used to describe an entity x is acceptable only if S stratifies x with respect to M and a granularity level g specified by the construction (formally: $\text{SR}_{M,g}(S)(x)$).

This constraint makes reference to the granularity level g , which different distributive constructions set to different values. Distributivity in *each*-constructions is always over atoms, while distributivity in *for*-adverbials and pseudopartitives is over parts which are smaller than the whole entity or event when measured along the dimension μ .

(69) **Granularity levels of different distributive constructions**

a. *Each*-constructions: $\lambda x. \text{Atom}(x)$, or more simply, Atom

b. Temporal *for*-adverbials: $\lambda t. t < \tau(e)$, where e is the event described by the *for*-adverbial.

- c. Pseudopartitives: $\lambda d.d < \mu(x)$, where μ is the Map and x is the entity described by the pseudopartitive.

The granularity parameters at work in (69b) and (69c) can be read as “the set of time intervals that are properly contained in the runtime of e ” and “the set of degrees that are smaller than the μ of x .” Let me define a helper function γ that abstracts over the pattern in these parameters. Here, α and β are any basic types, and $<$ abstracts over the mereological proper-part relation and the smaller-than relation over degrees.

$$(70) \quad \gamma \stackrel{\text{def}}{=} \lambda M_{(\alpha, \beta)} \lambda x_{\alpha} \lambda d_{\beta}. d < M(x)$$

From now on, using this function, I will abbreviate $\lambda t.t < \tau(e)$ as $\gamma(\tau, e)$, and I will abbreviate $\lambda d.d < \mu(x)$ as $\gamma(\mu, x)$. The predicates $\gamma(\tau, e)$ and $\gamma(\mu, x)$ have divisive reference (see Section 2.3.5), and so does (vacuously) the predicate $\lambda x.\text{Atom}(x)$. This reflects the intuition that granularity parameters are thresholds that exclude anything below them from consideration.

4.7 Compositional implementation

The Distributivity Constraint does not specify anything about compositional implementations. This section shows one possible way to incorporate it into a compositional framework, namely by using the LFs presented in Chapter 3.

I prefer to remain neutral on the status of violations of the Distributivity Constraint, for example regarding the question whether they represent presupposition failures or violations of a backgrounded part of the truth conditions. This question is orthogonal to my main claims. For the purposes of implementational concreteness only, I implement the constraint as a selectional restriction on the words that occur necessarily as part of a distributive construction. Following McCawley (1968) and Singh (2007), I assume that selectional restrictions are a kind of lexical presupposition.

For example, in *for*-adverbials, the presupposition is attached to the word *for*; in *each*-constructions, to the word *each*; and in pseudopartitives, to the word *of*. In a framework that uses partial truth-values in the style of Karttunen & Peters (1979) or Muskens (1996), a lexical presupposition is expressed as a restriction on the possible input values of a function. I write $\lambda x : \varphi . \psi$ for the partial function that is defined for all x such that its lexical presupposition φ holds, and that returns ψ wherever the function is defined. Sentences are interpreted as pairs of propositions: an assertion and a presupposition. These presuppositions must be fulfilled in order for the sentence to have a truth-value. I assume that the global presupposition of a sentence is the conjunction of all the lexical presuppositions associated with its lexical items. That is, all presuppositions project straight to the top. Sentences whose global presupposition is true have the same truth-value as their assertion; sentences where it is false lack a truth-value. Denotations of lexical items that carry a presupposition are

represented as partial functions that are undefined whenever this lexical presupposition is false.

The following lexical entries all embody the Distributivity Constraint. They are identical to the skeletal entries in Chapter 3, except that I have added a lexical presupposition. The letters K, S, and M stand for Key, Share, and Map respectively.

$$(71) \llbracket \text{of} \rrbracket = \lambda S_{\langle \alpha, t \rangle} \lambda M_{\langle \alpha, \beta \rangle} \lambda K_{\langle \beta, t \rangle} \lambda x_{\alpha} : SR_{M, \gamma(M, x)}(S)(x) . S(x) \wedge K(M(x))$$

$$(72) \llbracket \text{for} \rrbracket = \lambda K_{\langle i, t \rangle} \lambda M_{\langle v, i \rangle} \lambda S_{\langle v, t \rangle} \lambda e : SR_{M, \gamma(M, e)}(S)(e) . \\ S(e) \wedge K(M(e)) \wedge \text{regular}(M(e))$$

$$(73) \llbracket \text{each} \rrbracket = \lambda S_{\langle v, t \rangle} \lambda M_{\langle v, e \rangle} \lambda K_{\langle e, t \rangle} \lambda e : SR_{M, Atom}(S)(e) . S(e) \wedge K(M(e))$$

The skeletal LFs from Chapter 3 are repeated with these entries added in Figures 4.1 through 4.4 at the end of this chapter. These entries and full LFs are my official proposal for *for*-adverbials and pseudopartitives; the entry for *each* will be further revised in Chapter 9. The types of these entries reflect the specific order in which they combine with their constituents according to these LFs. This order is not essential and could be easily changed without consequences for the theory.

For ease of reference, here is the translation of a temporal *for*-adverbial:

$$(74) \llbracket \text{for an hour} \rrbracket \\ = \lambda P_{\langle v, t \rangle} \lambda e : SR_{\tau, \gamma(\tau, e)}(P)(e). \\ P(e) \wedge \text{hours}(\tau(e)) = 1 \wedge \text{regular}(\tau(e))$$

This translation is obtained as a result of combining the entry for *for* in (72) with *an hour* and with a [runtime] head as shown in Figure 4.3.

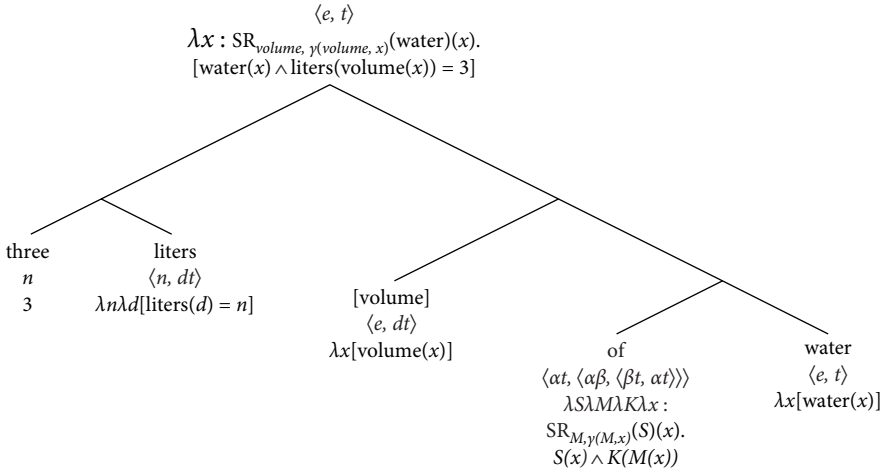


FIG. 4.1 Full LF of an ordinary pseudopartitive.

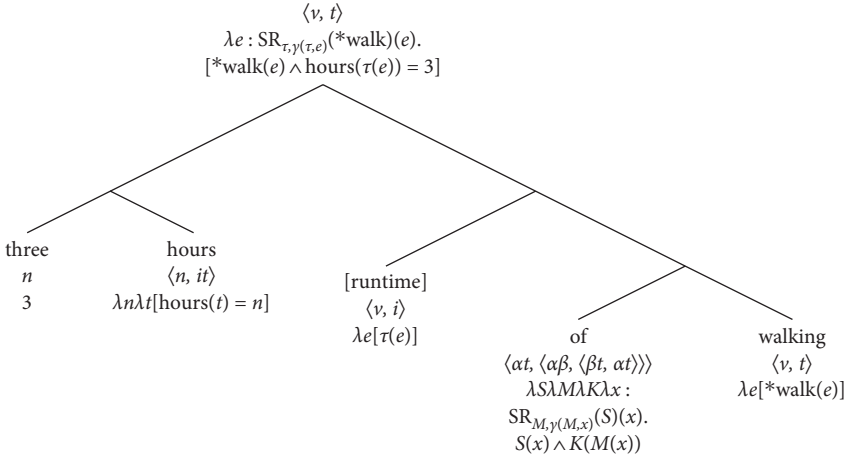
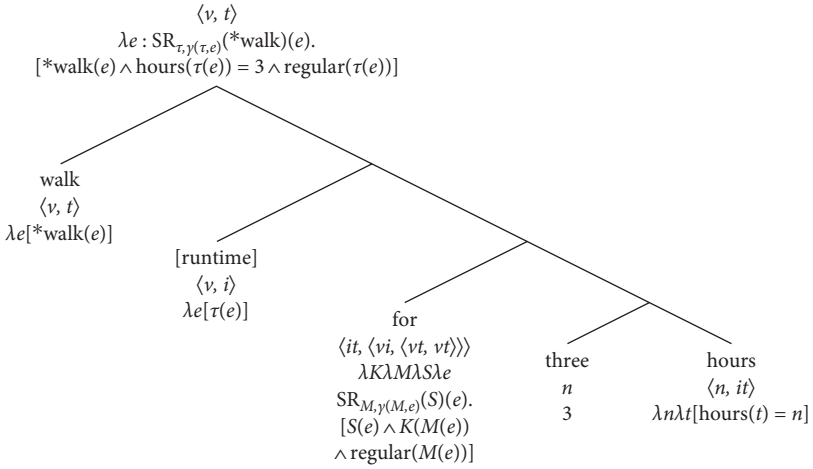


FIG. 4.2 Full LF of an event pseudopartitive.

FIG. 4.3 Full LF of a *for*-adverbial.

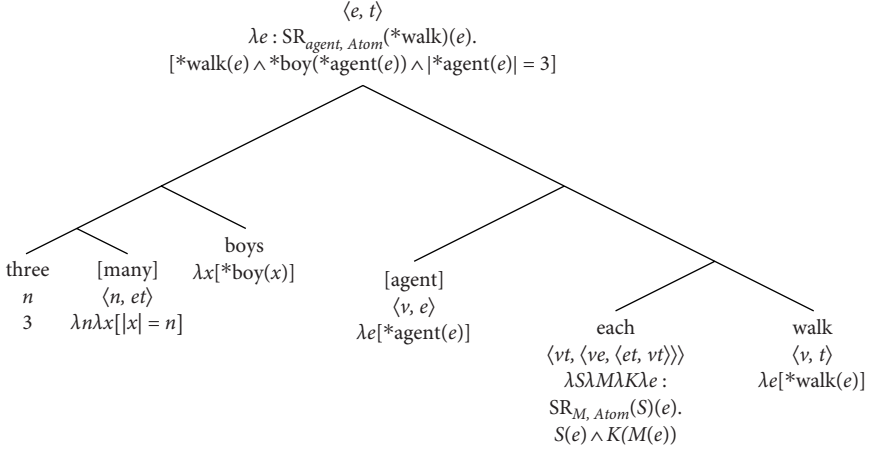
Also for ease of reference, I spell out and expand the results of the computations in Figures 4.1 through 4.4.

(75) $\llbracket \text{three liters of water} \rrbracket =$

$\lambda x : \text{SR}_{\text{volume}, \gamma(\text{volume}, x)}(\text{water})(x).$

$[\text{water}(x) \wedge \text{liters}(\text{volume}(x)) = 3]$

(For any entity x , this function is defined iff *water* stratifies x with dimension *volume* and granularity $\gamma(\mu, x)$. If defined, it is true iff x is a water amount whose volume measures three liters.)


 FIG. 4.4 Full LF of an adverbial-*each* construction.

- (76) $\llbracket \text{three hours of walking} \rrbracket =$
 $\llbracket \text{walk for three hours} \rrbracket =$
 $\lambda e : \text{SR}_{\tau, \gamma(\tau, e)}(*\text{walk})(e).$
 $[*\text{walk}(e) \wedge \text{hours}(\tau(e)) = 3 \wedge \text{regular}(\tau(e))]$
 (For any event e , this function is defined iff *walk* stratifies e with dimension τ and granularity $\gamma(\tau, e)$. If defined, it is true iff e is a regular walking event whose runtime measures three hours.)
- (77) $\llbracket \text{three boys each walked} \rrbracket =$
 $\lambda e : \text{SR}_{\text{agent}, \text{Atom}}(*\text{walk})(e).$
 $[*\text{walk}(e) \wedge *\text{boy}(*\text{agent}(e)) \wedge |*\text{agent}(e)| = 3]$
 (For any event e , this function is defined iff *walk* stratifies e with dimension *agent* and granularity *Atom*. If defined, it is true iff e is a walking event whose plural agent is three boys.)

The definedness conditions of these functions expand as in (78) through (80). I assume that these definedness conditions are all fulfilled as a matter of world knowledge or lexical semantics. The expansions in (78) through (80) can be understood as entailments of meaning postulates. Their paraphrases should make it clear that this is a plausible assumption.

- (78) $\text{SR}_{\text{volume}, \gamma(\text{volume}, x)}(\text{water})(x)$
 $x \in * \lambda y \left(\begin{array}{l} \text{water}(y) \wedge \\ \text{volume}(y) < \text{volume}(x) \end{array} \right)$
 (The water amount x in question can be exhaustively divided into parts (“strata”) which are water amounts and whose volumes are smaller than the volume of x .)

(79) $SR_{\tau, \gamma(\tau, e)}(*walk)(e)$

$$e \in * \lambda e' \left(\begin{array}{l} *walk(e') \wedge \\ \tau(e') < \tau(e) \end{array} \right)$$

(The walking event e in question can be exhaustively divided into parts (“strata”) which are also walking events and whose runtimes are properly included in the runtime of e .)

(80) $SR_{agent, Atom}(*walk)(e)$

$$e \in * \lambda e' \left(\begin{array}{l} *walk(e') \wedge \\ Atom(agent(e')) \end{array} \right)$$

(The walking event e in question can be exhaustively divided into parts (“strata”) which are also walking events and whose agents are atoms.)

4.8 Summary

Chapter 1 has presented the parallelism between the telic/atelic, collective/distributive, singular/plural, and count/mass oppositions in terms of boundedness. Intuitively, singular, telic, and collective predicates are delimited or bounded in ways that plural, mass, atelic, and distributive predicates are not. I pointed out that making formal sense of this parallelism amounts to answering what I called the *boundedness question*: How can the difference between boundedness and unboundedness be formally characterized?

This chapter has presented my answer to the boundedness question. The higher-order property of stratified reference characterizes what it means to be unbounded. This property is parametrized, which reflects the fact that unboundedness can be understood in more than one way. As the example *run to the store* shows, one and the same verb phrase can be distributive (unbounded with respect to agents) and telic (bounded with respect to runtime). As the example *kill* shows, one and the same verb can be collective on its agent position (bounded with respect to agents) and distributive on its theme position (unbounded with respect to themes). As the example *cable* shows, one and the same predicate can be divisive along its length (bounded with respect to length) and nondivisive along its diameter (unbounded with respect to diameter).

The picture presented in this chapter is idealized in several respects. For example, I have not discussed the fact that distributive entailments do not always literally hold as far down as they can. *For*-adverbials are compatible with predicates like *waltz* and *sleep in the attic* that do not satisfy the subinterval property (Dowty 1979). Similarly, pseudopartitives are compatible with heterogeneous mass nouns like *fruit cake* (Taylor 1977). The next chapter motivates the granularity parameter of stratified reference. As we will see, this parameter “makes room” for such predicates in distributive constructions.

Minimal parts

5.1 Introduction

In strata theory, as presented in Chapter 4, the Shares of adverbial-*each* constructions, *for*-adverbials, and pseudopartitives are characterized with the help of a novel higher-order property called stratified reference. The Shares of these constructions are distributive predicates, atelic predicates, and mass or plural predicates, respectively. With respect to adverbial-*each* constructions, I have argued that stratified reference is an adequate way to transfer traditional notions of predicative distributivity into an event-based framework. With respect to *for*-adverbials and pseudopartitives, I have briefly argued that stratified reference surpasses the subinterval property and divisive reference, which are the properties commonly used for the purpose of characterizing atelic, plural, and mass reference. In this chapter, I present the case against the subinterval property and against divisive reference in more detail. The case is based on what has been known at least since Dowty (1979) as the minimal-parts problem: entities and events that satisfy an atelic or mass predicate may have parts that are too small to still satisfy the predicate. I compare my implementation based on stratified reference with previous accounts of the minimal-parts problem (Dowty 1979, Hinrichs 1985, Moltmann 1989, 1991, Link 1991b). I argue that these previous accounts suffer from various technical deficiencies, and that stratified reference captures the essence of the minimal-parts problem in a more accurate way.

Section 5.2 discusses the minimal-parts problem in more detail, and Section 5.3 reviews some previous approaches to it. Section 5.4 shows how the granularity parameter of stratified reference solves the minimal-parts problem and explores its connection with the concept of cover. Section 5.5 concludes.

5.2 The minimal-parts problem

The minimal-parts problem occurs in the domains of verbal and nominal predicates (Taylor 1977, Dowty 1979). I discuss each of them in turn.

5.2.1 *The minimal-parts problem for verbal predicates*

As discussed in Section 4.5.3, the entailment properties of atelic predicates with respect to a large range of phenomena can be represented as a first approximation in terms of the subinterval property (Bennett & Partee 1972, Dowty 1979). A predicate *P* has the subinterval property iff whenever it holds at an interval, it also holds at every one of its subintervals. The minimal-parts problem consists in the fact that the subinterval property applies to certain atelic predicates such as *chuckle* (Taylor 1977) and *waltz* (Dowty 1979) only if we are willing to disregard subintervals below a certain threshold. For example, waltzing takes at least three steps, and there is no consensus on whether or not a given event whose runtime is substantially shorter than this three-step threshold still qualifies as waltzing. For this reason, it is difficult to know whether the subinterval property applies to *waltz*, and it is not clear whether this property is really an adequate characterization of atelicity. We can perhaps already conclude from this fact that the subinterval property must be given up. This conclusion is sometimes objected to on the grounds that the problem is not part of linguistics and should be ignored. Such an objection assumes that we do not really know what kinds of events occur at instants or even at very short subintervals, or in any case that these events should be seen as irrelevant for the purposes of semantics because they do not enter our daily experience. Once such events are excluded, the subinterval property can still technically be assumed to hold of all atelic predicates.

I think this kind of objection misses the point. There is no benefit in choosing the underlying assumptions of semantic theory with the sole purpose of being able to continue using the subinterval property, because one might as well avoid the problem by using a different property than the subinterval property to begin with. The extent to which our intuitions about short events would have to be changed in order to maintain that atelic predicates have the subinterval property is considerable. For example, a minimal event in the denotation of the predicate *pass on from generation to generation* will span several generations (Zwarts 2013, Doetjes 2015). In order to give this predicate in (1) the subinterval property, even sequences of several generations would have to be considered as having minimal time:

- (1) The Chinese people have created abundant folk arts, such as paper-cuttings, acrobatics, etc., passed on from generation to generation for thousands of years.¹⁹

It does not seem practical to maintain that the predicate *pass on from generation to generation* really has the subinterval property. But the predicate is clearly atelic since it can be modified by a *for*-adverbial. I conclude that atelic predicates do not necessarily have the subinterval property. This conclusion is not new. As we will see,

¹⁹ Attested example, accessed May 25, 2016. <http://web.archive.org/web/20131012194104/http://www.twinbridge.com/detail.aspx?ID=315>.

Dowty (1979) already recognized that the subinterval property is an idealization, and many subsequent authors have avoided using this property in their formalizations (see Section 5.3).

5.2.2 The minimal-parts problem for mass nouns

An analogous concept to the subinterval property, divisive reference, is sometimes used following Cheng (1973) as a defining semantic property of mass nouns: any part of something denoted by a mass noun is assumed to be denoted by the same mass noun. The definition of divisive reference is repeated here from Section 2.3.5:

(2) **Definition: Divisive reference**

$$\text{DIV}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$$

(A predicate P is divisive iff whenever it holds of something, it also holds of each of its proper parts.)

The assumption that divisive reference holds of mass denotations runs into a well-known problem with very small parts of mass substances. This problem is analogous to the one that arises in connection with the subinterval property, and concerns the atomic nature of matter: arguably hydrogen atoms do not qualify as water, so if we consider hydrogen atoms to be part of water, then not every part of water is water (Quine 1960). A different form in which this problem occurs comes from heterogeneous mass predicates like *fruit cake*, *pea soup*, or *succotash* (e.g. Taylor 1977): for example, a portion of fruit cake may contain sultanas, but these sultanas do not themselves qualify as fruit cake. This type of argument is arguably stronger than Quine's argument, because in this case, the minimal parts that do not qualify are directly accessible to the senses. While it might be possible to save the assumption that mass nouns have divisive reference by assuming that we conceptualize even predicates like *fruit cake* as having divisive reference, this assumption is even more difficult to justify in the case of fake mass nouns like *furniture*. Psycholinguistic experiments suggest that the cognitive structures underlying fake mass nouns are more similar to those of count nouns than to those of other mass nouns (see Section 2.4.1). It would be implausible both in light of the factual reality and in light of our cognitive modeling of it to assume that fake mass nouns have divisive reference (Barner & Snedeker 2005, Chierchia 2010).

The conclusion that mass nouns do not always have divisive reference appears to be generally accepted. Gillon (1992: 598) notes that “[w]hile some semanticists retain the divisivity of reference as a criterion to distinguish mass nouns from count nouns . . . only Bunt (1979, 1985) has attempted to justify the retention.” Gillon gives counterarguments to Bunt's reasons and argues that the grammar is mute on whether or not mass terms have divisive reference. This is also the position I adopt here.

5.3 Previous accounts of the minimal-parts problem

The significance of the minimal-parts problem lies in the challenges it poses for formalizing atelicity checks and checks for mass reference. For example, a proposal that explains why *for*-adverbials reject telic predicates by assuming that *for*-adverbials check for the subinterval property faces the minimal-parts problem. Solutions to the problem consist of reformulations of the subinterval property in a way that retains the ability to distinguish between atelic and telic predicates while correctly categorizing predicates like *waltz*.

This section reviews a representative sample of previous approaches to the minimal-parts problem that gives an idea of the space of possibilities. I discuss the promising aspects and the technical problems of these proposals. I have not attempted to include every approach known to me. For example, I do not discuss the interesting accounts in Vlach (1993), Piñón (1999a), van Geenhoven (2004), or Landman & Rothstein (2009, 2012a, 2012b). Other reviews of approaches to the minimal-parts problem are found in Krifka (1989b: ch. 2) and Mollá-Aliod (1997: ch. 4). I concentrate on the domain of aspect rather than on mass reference, because this is where most solutions seem to have been proposed.

5.3.1 Dowty (1979)

The first formal proposal for a model-theoretic translation of *for*-adverbials is found in Dowty (1979: 333). His analysis is prefigured in Dowty (1972). Dowty models *for*-adverbials as universal quantifiers ranging over the subintervals of some time interval. In other words, he essentially implements the subinterval property directly, and as a part of the asserted content rather than as a presupposition. Here is Dowty's proposal in its original form.

- (3) a. $\llbracket \text{for} \rrbracket = \lambda P_t \lambda P \lambda x [P_t\{n\} \wedge \forall t[t \subseteq n \rightarrow \text{AT}(t, P\{x\})]]$
 b. $\llbracket \text{an hour} \rrbracket = \lambda t[\text{an-hour}(t)]$

This entry is used in an intensional system where formulas are interpreted relative to a world-time index $\langle w, i \rangle$. The notation $P\{x\}$ is an abbreviation for $\mathcal{P}(x)$, the result of applying the extension of P to x . Under any world-time index $\langle w, i \rangle$, the indexical constant n denotes the time index i . Phrases like *an hour* and *six weeks* are interpreted as predicates of times. For example, *an hour* denotes the set of all contiguous temporal intervals whose duration is an hour. $\text{AT}(t, p)$ evaluates a proposition p at a time index t .

In order to translate the sentence *John slept for an hour*, Dowty combines *for* with *an hour*, *sleep*, *john*, and with an operation that introduces an existential quantifier over past times. I abstract away from various details of his analysis not relevant here. In particular, his analysis is embedded in an intensional system, but this fact plays no role in his analysis of *for*-adverbials. Without its irrelevant parts, Dowty's translation

of this sentence is represented in simplified form in (4) (for the original analysis, see Dowty 1979: 334):

- (4) $\llbracket \text{John slept for an hour} \rrbracket$
 $= \exists t_1 [\text{an-hour}(t_1) \wedge \forall t_2 [t_2 \subseteq t_1 \rightarrow \text{sleep}(j, t_2)]]$
 (There is a time interval t_1 which lasts an hour and John sleeps at each of its subintervals.)

The representation in (4) essentially implements the subinterval property one-to-one, because it quantifies over all subintervals of the interval t_1 . Dowty is aware of the minimal-parts problem and suggests that to account for it, quantification should not be over literally all subintervals, as in this representation, but only over “all subintervals large enough to be minimal intervals for the activity in question”; he notes that how to do this is unclear. For simplicity, he decides to leave his analysis as it stands.

For many purposes, Dowty’s analysis of aspect is adequate even though (or because) it ignores the minimal-parts problem. It is adopted by many authors who are concerned with the interaction of *for*-adverbials with other semantic components, in particular the Perfect, rather than with the meaning of *for*-adverbials themselves. This line of work includes Richards (1982), Heny (1982), Mittwoch (1988), Parsons (1990), Abusch & Rooth (1990), Kamp & Reyle (1993), Hitzeman (1997), Iatridou, Anagnostopoulou & Izvorski (2001), and Rathert (2004). The latter contains an excellent review of this part of the literature. Here, I ignore these approaches and focus on the minimal-parts problem.

5.3.2 Hinrichs (1985)

Hinrichs (1985), a dissertation supervised by Dowty, assumes a Davidsonian event-based semantics. In order to avoid the minimal-parts problem, Hinrichs relaxes the requirement that *for*-adverbials place on their predicates. Hinrichs’ proposal is complicated by the fact that he adopts the three-tiered ontology of kinds, objects, and stages, which Carlson (1977) proposed in order to account for the properties of bare plurals and generic sentences. Abstracting away from this and other complexities, here is how *John slept for an hour* comes out on his proposal:

- (5) $\llbracket \text{John slept for an hour} \rrbracket$
 $= \exists e \exists l [\text{hour}(l) \wedge l \leq \tau(e) \wedge \text{sleep}(j)(e)] \wedge$
 $\forall l' [l' < l \rightarrow \exists e' [e' < e \wedge l' \leq \tau(e') \wedge \text{sleep}(j)(e')]]$

Hinrichs comments on his translation as follows (1985: 235):

The translation in [(5)] requires that . . . there has to be a spatio-temporal location l with the property of being one hour long such that the entire process of John’s sleeping spatio-temporally contains l and for each proper sublocation l' of l there has to be a proper subprocess of John’s sleeping containing l' [W]e don’t require that each sublocation denoted by the temporal

measurement phrase has to *be* a subprocess itself. The requirement that each proper sublocation be contained in a proper subprocess of the maximal process making up the event, rules out that for each sublocation we could simply pick the maximal process itself.

The way Hinrichs avoids the minimal-parts problem does not follow Dowty's suggestion of quantifying over less than literally all subintervals. Instead of requiring that a predicate like *John sleep* be true at each or most subintervals, Hinrichs requires that the runtime of every subinterval of John's sleeping must be contained in (and not necessarily equal to) that of a sleeping event. This sleeping event must be a proper part of the event that the sentence describes. This is possible because *sleep* is not quantized; if we assume that all telic predicates are quantized, they are correctly ruled out. The assumption that all telic predicates are quantized is not available to us because it is incompatible with the assumption of lexical cumulativity (see Section 2.7.2). But Hinrichs (1985) does not assume lexical cumulativity.

Krifka (1989b: 150) criticizes Hinrichs' approach as arbitrary and notes that the proper-part requirement does not explain anything, since it does not serve any purpose other than that of excluding telic predicates. There is also another problem for Hinrichs' account, which causes it to run into something very similar to the minimal-parts problem. The variable l' in (5) ranges over subintervals rather than (just) over instants. On the assumption that time is dense, it ranges over subintervals that are only minimally shorter than l . For example, it ranges over intervals of length 58, 59, $59\frac{1}{2}$. . . minutes. For each of these intervals, the definition requires there to be a proper part of the event of John's sleeping which lasts at least as long as the interval and which qualifies as John's sleeping. This gives the events in the denotation of atelic predicates a dense structure. Moreover, depending on the structure of the underlying mereology, there will generally be a unique complement that results from removing the smaller sleeping event from the larger sleeping event. In classical extensional mereology (CEM), this is a consequence of Unique Separation (see Section 2.3.2). Depending on how large a sleeping event one separates off, the runtime of its complement may become arbitrarily small. It follows on Hinrichs' account that only events with arbitrarily short parts can be described by a sentence with a *for*-adverbial. This seems like an unintended consequence of the account, and it commits us to the assumption that no events in the extension of atelic predicates can ever be mereological atoms. My own proposal avoids making this assumption.

At first sight it might look like Hinrichs' approach could be salvaged by restricting l' to range only over instants and not intervals. However, this requires the assumption that time is atomic, otherwise there are no instants. According to von Stechow (2009), this assumption is rejected by most semanticists, and I do not rely on it myself (see Section 2.4.4).

The analysis of Hinrichs (1985) is used in slightly adapted form by Abusch & Rooth (1990), but their adaptations do not prevent the problem from arising. A more

significant reformulation of Hinrichs' proposal is found in Rathert (2004), where it is rendered using the following representation:

- (6) $\llbracket \text{John ran for two hours} \rrbracket(t) = 1$
 iff $|t| = 2 \text{ hours} \wedge \forall m \subseteq t \exists n[m \subseteq n \subseteq t \wedge \llbracket \text{John ran} \rrbracket(n) = 1]$, with n being a minimal run-event (two steps, done faster than a walk)

This reformulation runs into problems of its own: for example, since m is not required to be an instant, it ranges over some subintervals which are too big to fit into the runtimes of any *minimal* run-events. In fairness, Rathert is less interested in the meaning of *for*-adverbials than in how they interact with tense and with the Perfect. Still, we see that the proposal of Hinrichs (1985) is technically flawed and has resisted several attempts to repair it.

5.3.3 Moltmann (1989, 1991)

Moltmann (1989, 1991) recasts Dowty's analysis in an event-semantic framework, and follows it quite closely. Her representation of an atelic predicate is as given in (7):

- (7) $\llbracket \text{John slept for an hour} \rrbracket$
 $= \exists t_1[\text{an-hour}(t_1) \wedge \forall t_2[t_2 P t_1 \rightarrow \exists e. \text{sleep}(e, \text{john}) \wedge \tau(e) = t_1]]$
 "There is a time interval t_1 which lasts an hour and John sleeps at each of its relevant parts."

Where Dowty has the relation \subseteq and where the present framework would use the mereological relation \leq , Moltmann uses a relation she calls P . The most recent presentation of the formal model that describes the properties of this relation P is found in Moltmann (1997, 1998). This framework departs in significant ways from the basic axioms of CEM (see Section 2.3): the relation P is not assumed to satisfy transitivity, and the domains of individuals and events are not assumed to be closed under sum (see also my brief discussion of her views in Section 2.3.1). With respect to *for*-adverbials, Moltmann (1991) summarizes the relevant aspects of her model as follows:

"The part structure of an interval cannot be taken as being strictly divisive in a mathematical or physical sense. Rather, it appears that semantics involves a coarser part structure and a notion of relevant or contextually determined part, namely the relation P . Depending on the type of event, the part structure of the interval must have smallest subintervals of a certain minimal length. This is required, for instance, when the event is a process such as writing (not any physical part of a writing event is considered as writing) or a repetitive event (not any part of a repetitive revolving is a revolving). Therefore, the intended meaning of ' P ' is the relation 'is a relevant part of', a relation which does not involve any subinterval of the measuring interval." (Moltmann 1991: 638–9)

" P has to be understood not as a part relation in a strict mereological sense, but rather as a contextually determined relation that may be coarser than the mereological part relation, as

the relation ‘is relevant part of’ . . . One and the same entity may have different part structures depending on the respective context. For instance, a time interval may be conceived of as consisting of smallest subintervals of different length in different contexts—depending, for instance, on the type of events that are under consideration.” (Moltmann 1991: 633)

Moltmann’s departure from CEM seems to me too radical a step, especially because the formal system by which she replaces it is not nearly as constrained and well understood as CEM. This makes it difficult to evaluate the predictions of her theory other than those intended by her. As Zucchi & White (2001: 233, n. 5) put it: “Since Moltmann does not tell us much about what *relevant* parts are, it is unclear to what extent her formulation actually solves the minimal parts problem.” A more thorough criticism of Moltmann’s position from the point of mereology is found in Pianesi (2002). See also Varzi (2006) for related discussion.

One possible way to make sense of Moltmann’s suggestions could be the proposal in Link (1987b), which sketches a way to resolve the minimal-parts problem by replacing the lattice of events with a partial order of lattices E_i , with each of them standing for a certain granularity of events: intuitively, if $j \leq k$ then the events in E_j are more fine-grained than the events in E_k . A family of homomorphisms h_{ij} between the lattices relates events to events in such a way that the part relation between these events is partially collapsed in coarser models but otherwise preserved. Formally, h_{ii} is the identity map on E_i , and for any three lattices E_i, E_j, E_k such that $i \leq j \leq k$, the map is constrained by $h_{ik} = h_{ij} \circ h_{jk}$. This allows for the possibility that “a certain event can be atomic in a coarser domain and still be the image under a h_{ij} of a complex sum of events in a more fine-grained E_j .” Link proposes to treat the granularity of events as a discourse parameter.

Unfortunately, Link’s proposal is difficult to evaluate because he does not go into more detail than that. It seems to me that a great deal of complexity is added to the model by appealing to a family of lattices. A later version of his proposal is described in Section 5.3.4.

5.3.4 Link (1991b)

Link (1991b) does not deal with *for*-adverbials per se, but he proposes a modification of the subinterval property that excludes from consideration temporal intervals below a certain threshold. He calls this threshold the *granularity parameter*. The modified subinterval property is designed to be able to hold of activity predicates despite the minimal-parts problem.

Link’s formulation of the subinterval property is shown in (8):

- (8) $\text{SUB}_{\text{Link}}(P) \stackrel{\text{def}}{=} \forall e, t [P(e) \wedge t \leq \tau(e) \wedge |t| \geq \gamma(e) \rightarrow \exists e' (e' \leq e \wedge P(e') \wedge \tau(e') = t)]$
 (A predicate P has the **Link-style subinterval property** iff for any e of which it holds, every part of the runtime of e whose length is at least $\gamma(e)$ is the runtime of a part of e of which P holds as well.)

Link calls $\gamma(e)$ the *granularity* of e and comments that this parameter “expresses the observation that some time stretches might be too small to be a trace of the event e ; thus $\gamma(e)$ fixes the minimal length that a time stretch must have to serve as a trace for e ” (p. 217). Link’s implementation therefore quantifies over all temporal intervals whose length is $\gamma(e)$ or larger. This implementation essentially requires every partition of the runtime of an event into minimal-size subintervals to be a set of runtimes of P -events.

While the granularity intuition is on the right track, this condition seems too strict. Take a sentence such as this one, based on Link (2015):

(9) John swam laps for an hour.

Suppose that John swims sixty laps in succession starting at 1pm, one lap per minute. Let e_o be the sum of these events. Then (9) is true of e_o . We would want to express that the predicate $P = \text{swim laps}$ has the Link-style subinterval property with granularity one minute.

But this is not possible in Link’s implementation (see also Link 2015 for a similar observation). We would expect P to have the subinterval property with respect to e_o and we set $\gamma(e_o) = 1\text{min}$. Since e_o has a runtime i that lasts from 1pm to 2pm, Link’s definition requires every subinterval of i that lasts one minute or more to be the runtime of an event in which John swims laps. The problem is that not every subinterval of i satisfies this condition. If John starts each lap on the full minute, there will be no lap from 1:00:30 pm to 1:01:30 pm, and so there will be no event at that time that qualifies as *swim laps*, not even if *laps* is interpreted inclusively as denoting *one or more laps* (see Section 2.6.2).

5.3.5 Borik (2006)

Given the problems faced by accounts that are based on universal quantification over subevents, one might wonder if existential quantification over subevents might be more appropriate. One account that adopts this stance is Borik (2006: 53f.). According to Borik, a predicate P is *atelic* iff every interval i at which it holds has a (proper) subinterval i' at which P also holds, given that the denotations of all the nominal arguments of P remain the same. Intuitively, this means that atelicity is defined as being nonquantized along the temporal dimension.

The examples in (10) show that this requirement is not adequate (see Krifka 1992):

- (10) a. *The courier ran all the way to Athens for three hours.
b. The courier ran towards Athens for three hours.

Suppose the predicate *The courier ran all the way to Athens* holds at a three-hour interval i . Since it only mentions the endpoint but not the origin of the run, it will also hold at any final subinterval of i . The only possible exception concerns very short subintervals of i because they may be too short to qualify even if their endpoint is Athens; this is the minimal-parts problem. But we cannot rely on the minimal-parts

problem to rule out (10a) without also ruling out (10b). I conclude that existential quantification over proper subintervals is too weak as a formalization of atelicity.

5.3.6 *Summing up*

The account by Dowty (1979) ignores the minimal-parts problem, though Dowty himself is aware of it; the account by Hinrichs (1985) seeks to avoid it by a mathematical trick but unintentionally reintroduces a related problem; Moltmann (1991) throws the baby out with the bathwater (at least in my eyes) by giving up mereology in favor of a very minimal theory of part structures; the account by Link (1991b) is too strict because it requires the main-clause predicate to be true at every subinterval of a certain length; and the account by Borik (2006), which replaces universal by existential quantification over subintervals, imposes a requirement that is too weak because it fails to diagnose telicity in motion predicates that only specify their endpoint.

In addition to these technical problems, most of these accounts except for Moltmann's are limited in scope because they restrict themselves to atelic predicates, which is just one place where the minimal-parts problem arises. This exposes them at least in principle to the problem of being descriptions rather than explanations. Here I follow a simple concept of explanation suggested in von Stechow (1984b) for the purposes of comparing semantic theories: "If a number of highly complex and apparently unrelated facts are reducible to a few simple principles, then these principles explain these facts."²⁰ Section 5.4 motivates my formulation of stratified reference, which was already presented in Chapter 4, and shows that this formulation avoids the minimal-parts problem.

5.4 Stratified reference as a solution to the minimal-parts problem

In this section, I further motivate the formulation of stratified reference introduced in Chapter 4 and discuss its connection with the concept of cover.

Many of the solutions to the minimal-parts problem in Section 5.3 share the intuition that one must carve out an exception for very small entities in the higher-order property that is used to characterize atelicity and mass reference. For example, instead of quantifying over literally all subintervals as the subinterval property does, the solution by Link (1991b) avoids quantifying over certain subintervals which are below a certain granularity threshold. The discussion of Link's implementation in Section 5.3.4 has shown that it will not do to simply fix a threshold, exclude all of the subintervals below it, and quantify over all the subintervals above it. A sentence like *John swam laps for an hour* does not require every sufficiently long subinterval of

²⁰ According to von Stechow, this concept of explanation is also present in syntactic work of generative grammar in the Chomskian tradition, but he cites Chomsky (1982) as rejecting it for semantics.

a one-hour interval to be the runtime of a lap-swimming event. Rather, this sentence is true in virtue of the fact that this one-hour interval can be divided into subintervals at each of which John swims a lap. The minimal-parts problem then consists in specifying the right duration for these subintervals in such a way as to make sure that each subinterval has to be shorter than the whole interval without having to be arbitrarily short. In the present framework, sentences do not apply directly to intervals but to events, so we may instead say that the *for*-adverbial requires the sentence *John swam laps* to apply to an event that can be divided into subevents each of which is a lap-swimming event by John of the right duration. In mereology, dividing an event into a given set of subevents amounts to testing whether the event is the sum of these subevents. We may therefore equivalently say that the *for*-adverbial requires the sentence to apply to an event that is the sum of proper subevents which are each lap-swimming events by John of the right duration. The challenge then consists in specifying what that right duration amounts to.

To understand the formalization to be proposed, suppose for a moment that it is already known how to determine whether a subevent counts as having the right duration, and that there is a “cover predicate” C that encapsulates that knowledge. C is analogous to the granularity parameter γ in Link (1991b). We can now write $\lambda e'[C(e')]$ for the set of all subevents of the right duration, and $\lambda e'[P(e') \wedge C(e')]$ for the intersection of this set with a predicate P . Intuitively, P stands for the predicate whose telicity we want to test. Then the algebraic closure $*\lambda e'[P(e') \wedge C(e')]$ is the set of all sums of subevents of the right duration which are each in P . We can then describe the requirement that a *for*-adverbial imposes on a predicate P by stating that the predicate $*\lambda e'[P(e') \wedge C(e')]$ must apply to any event that the *for*-adverbial describes. This is expressed in the following preliminary definition:

(11) **Preliminary definition: Stratified subinterval reference**

$$\text{SSR}(P)(e) \stackrel{\text{def}}{=} e \in *\lambda e' \left(\begin{array}{c} P(e') \wedge \\ C(e') \end{array} \right)$$

(An event predicate P has stratified subinterval reference with respect to an event e iff e can be divided exhaustively into parts (“strata”) that are each in P and which satisfy the predicate C .)

I will update this definition in a moment by refining the unanalyzed predicate C . The result was already anticipated in Chapter 4. But first, let us strengthen the intuitions of what it means for an event to have a starred predicate like $*\lambda e'[P(e') \wedge C(e')]$ apply to it. For this purpose, it is useful to consider the relation between this kind of starred predicate and the concept of cover that figures prominently in the work of Gillon (1987) and Schwarzschild (1996) (as discussed in more detail in Section 8.3). In a set-based representation of plural individuals, covers are partitions of a set whose cells are allowed to overlap. For example, the set $\{\{a, b\}, \{b, c\}\}$ is a cover of the set $\{a, b, c\}$. This notion can be defined as follows:

(12) **Definition: Cover (set-theoretic)**

$$\text{Cov}(C, P) \stackrel{\text{def}}{=} \emptyset \notin C \wedge \bigcup C = P$$

(C is a cover of a set P iff C is a set of nonempty subsets of P whose union is P .)

As is often observed (e.g. Heim 1994), this notion can be translated into mereological terms. One possible way to do so is the definition in (13) (see also Section 2.3.2 for correspondences between set theory and mereology):

(13) **Definition: Cover (mereological)**

$$\text{Cov}(C, x) \stackrel{\text{def}}{=} x = \bigoplus C$$

(C is a cover of a mereological object x iff C is a set whose sum is x .)

As the following theorem illustrates, there is a close connection between the concept of cover and starred predicates:

(14) **Theorem:**

$$\forall x[x \in {}^*\lambda y[C(y)] \Leftrightarrow \exists C' \subseteq C[\text{Cov}(C', x)]]$$

(The algebraic closure of C applies to x iff there is a subset of C that is a cover of x .)

The proof of this theorem is simple. Using the definition of algebraic closure in Section 2.3.1, we rewrite $x \in {}^*\lambda y[C(y)]$ as $\exists C' \subseteq C[x = \bigoplus C']$. By definition (13), this is equivalent to $\exists C' \subseteq C[\text{Cov}(C', x)]$. For more discussion of this connection between algebraic closure and cover-based approaches, see Vaillette (2001).

Given this connection, my proposal amounts to describing the atelicity requirement of a *for*-adverbial in terms of a set C that must be a cover of any given event in the denotation of the predicate. Since C is a cover of these events, the cells of C must contain their parts. Given C , an atelic predicate can be defined as one such that each cell of C contains an event that is again in the denotation of the predicate.

So far, I have not said anything about the nature of this predicate C . We might suspect that C is set to some value that is very short in relation to the time interval denoted by the *for*-adverbial. I took this route in earlier versions of this theory (Champollion 2010b, 2015c) based on an observation by Moltmann (1991: n. 9) and related conjectures. According to Moltmann, given a context in which John draws two pictures per hour, a sentence like *For one hour John drew pictures* seems less acceptable than *For ten hours John drew pictures*. However, I have since then tried and failed to replicate this acceptability contrast via web surveys.²¹ Independently, Piñón (2015) and Schwarzschild (2015) point out that instead of using C to specify that the runtime of a subevent is sufficiently long, we may use it to specify that the runtime of a subevent is shorter than the runtime of the whole event. For more discussion, see

²¹ Surveys were run using TurkTools (Erlewine & Kotek 2016). I thank Hanna Muller and Linmin Zhang for their help with survey design and implementation.

Champollion (2015b). I will follow Piñón and Schwarzschild in assuming that all that is required in order to rule out telic predicates is given in (15), where e is the event that the *for*-adverbial describes:

$$(15) \quad C = \lambda t. t < \tau(e)$$

This predicate holds of intervals that are properly included in the runtime of e . In Section 4.5.4, I have introduced the shorthand $\gamma(\tau, e)$ for it. Given this, C contains any event whose runtime is properly included in the runtime of e , which is to say, the length of the interval expressed by the Key (the measure phrase of a *for*-adverbial). As a consequence, the size of the cells in C is determined differently from case to case. Depending on the sentence, C may contain only very short events or also quite long events. For a sentence like (1), repeated as (16), where the *for*-adverbial is *for thousands of years*, C will apply to some events whose runtime is barely shorter than a few thousand years.

- (16) The Chinese people have created abundant folk arts, such as paper-cuttings, acrobatics, etc., passed on from generation to generation for thousands of years.
= (1)

By contrast, for a sentence like (17), C will only apply to events whose runtime measures less than 3600 milliseconds:

- (17) Ded'leg says: How i [sic] stop a macro for 1sec? Cog says: By creating a script that will loop for 3600 [sic] milliseconds . . . depending on how long you think it will take for both your computer, the server and internet lag to affect the macro.²²

The updated definition of stratified subinterval reference, in which C is replaced by (15), is shown in (18):

- (18) **Definition: Stratified subinterval reference**

$$SSR(P)(e) \stackrel{\text{def}}{=} e \in * \lambda e' \left(\begin{array}{l} P(e') \wedge \\ \tau(e') < \tau(e) \end{array} \right)$$

(An event predicate P has stratified subinterval reference with respect to an event e iff e can be divided exhaustively into parts (“strata”) that are each in P and whose runtimes are properly included in the runtime of e .)

Section 4.5.3 has shown how this definition can be used to identify *waltz* as an atelic predicate, despite the minimal-parts problem. Section 4.6 has generalized the definition to stratified reference, repeated here as (19).

- (19) **Definition: Stratified reference**

Let d (a “dimension”) be any function from entities of type α to entities of type β ,

²² Attested example, accessed May 25, 2016. <http://web.archive.org/web/20120429171708/http://www.wowmacroswarcraft.com/faq/>.

and let g (a “granularity level”) be any predicate of entities of type β . Let P range over predicates of type $\langle \alpha, t \rangle$ where α is either e or v , and let x range over entities of type α . Then:

$$\text{SR}_{d,g}(P)(x) \stackrel{\text{def}}{=} x \in {}^*\lambda y \left(\begin{array}{c} P(y) \wedge \\ g(d(y)) \end{array} \right)$$

(A predicate P (the “Share”) stratifies x with respect to a function d (the “dimension” or “Map”) and a predicate g (the “granularity level”) iff x can be exhaustively divided into parts (“strata”) which are each in P and which are each mapped by d to something in g .)

Instead of testing whether the predicate to which it is applied holds at every subinterval, stratified reference—with the granularity parameter instantiated as in (15)—only tests whether the entity of which it holds can be divided into parts which are smaller than the entity in question. I call these entities *strata* as a reminder of the fact that they generally have small values as measured in one dimension, but may be arbitrarily large as measured in any other dimension. As discussed in Chapter 1, the term *strata* is chosen as an allusion to the geological formations which are visible in places such as the Grand Canyon. A geological stratum can be just a few inches thick and yet extend over hundreds of thousands of square miles. The dimension along which definition (60) constrains strata is specified by the parameter f ; their “maximum thickness” is specified via the threshold parameter g , which I have instantiated as in (15). This method avoids explicit reference to atomic events or intervals.

The novel connection that this account draws between *for*-adverbials and other distributive constructions makes it possible to formulate the solution to the minimal-parts problem in a completely general way that does not refer to *for*-adverbials specifically. The granularity parameter of stratified reference can be set to whatever value is desired. Here, I have chosen a value that essentially allows any predicate P to combine with a temporal *for*-adverbial as long as any event in P consists of at least two potentially overlapping temporally shorter parts that are each in P . A similar requirement was proposed in Kratzer (2007: 285), but without the requirement that the parts be temporally shorter.

This requirement is fairly weak. Of course, many atelic predicates P apply to events with many more than two P parts; however, this does not mean that *for*-adverbials require that many such parts exist. Intuitively, we may think of a *for*-adverbial as a sieve and of the events in the denotation of a predicate as grains of sand. Describing the requirement that *for*-adverbials impose corresponds to describing the size of the holes in the sieve, which may be bigger than the size of the grains that pass through it (Champollion 2015b).

A related but conceptually distinct task consists in describing what we know about the grain size of various atelic predicates. Some predicates like *run*, *loop*, and (to a

slightly lesser extent) *waltz* will be relatively fine-grained, while others like *pass on from generation to generation* will be more coarse-grained. Stratified reference can be used for this purpose as well. In Section 4.5.1, I have used meaning postulates to describe predicates that distribute down along various thematic roles. Likewise, we can formulate meaning postulates to describe predicates that distribute down along the temporal dimension. The postulate in (20) uses the universal version of stratified reference, defined in Section 4.6, to express that the shortest waltzing events are under three seconds long. This is an arbitrary value, and I use it only for concreteness.

(20) **Meaning postulate: *waltz* has stratified reference down to 3 seconds**

$$\text{SR}_{\tau, \lambda t[\text{seconds}(t) \leq 3]}(\lambda e. * \text{waltz}(e)) \Leftrightarrow$$

$$\forall e. * \text{waltz}(e) \rightarrow e \in * \lambda e' \left(\begin{array}{c} * \text{waltz}(e') \wedge \\ \text{seconds}(\tau(e')) \leq 3 \end{array} \right)$$

(Every waltzing event can be divided into one or more parts, each of which is a waltzing event whose runtime is at most three seconds long.)

For complex predicates like *pass on from generation to generation*, we may want to derive their grain size from the meanings of their parts rather than specifying it via meaning postulates. This is the purpose of theories of aspectual composition such as Krifka (1998). I will set this task aside here.

As discussed in Section 4.7, I assume that the temporal adverbial *for an hour* is translated as in (21):

(21) **[[for an hour]]** (my proposal, repeated from Section 4.7)

$$= \lambda P_{\langle v, t \rangle} \lambda e : \text{SR}_{\tau, \gamma(\tau, e)}(P)(e).$$

$$P(e) \wedge \text{hours}(\tau(e)) = 1 \wedge \text{regular}(\tau(e))$$

Given this translation, sentence (9) is assigned the following logical representation:

(22) **[[John swam laps for an hour]]**

$$\exists e : \text{SR}_{\tau, \gamma(\tau, e)}(\llbracket \text{swim laps} \rrbracket)(e).$$

$$[\llbracket \text{swim laps} \rrbracket(e) \wedge \text{hours}(\tau(e)) = 1 \wedge \text{regular}(\tau(e))]$$

(There is a regular event *e* in the denotation of *swim laps* whose runtime measures an hour, and it is presupposed that *swim laps* stratifies *e* with dimension τ and granularity $\gamma(\tau, e)$.)

Given the definition of stratified reference above, the presupposition of (22) expands as follows:

(23) $\text{SR}_{\tau, \gamma(\tau, e)}(\llbracket \text{swim laps} \rrbracket)(e) \Leftrightarrow$

$$e \in * \lambda e' \left(\begin{array}{c} \tau(e') < \tau(e) \wedge \\ \llbracket \text{swim laps} \rrbracket(e') \end{array} \right)$$

(The hour-long event *e* can be divided into parts each of which is in $\llbracket \text{swim laps} \rrbracket$ and has a runtime properly included in the runtime of *e*.)

This is an appropriately weak requirement. In the scenario discussed in Section 5.3.4, it is satisfied because *e* can be divided into sixty consecutive lap-swimming events. The fact that those subintervals of the runtime of *e* that do not start and end on the full minute are not runtimes of lap-swimming events does not preclude the requirement from being satisfied.

At the same time, the requirement is not so weak as to run into the problem faced by Borik (2006) in connection with the examples in (10), repeated here:

- (24) a. *The courier ran all the way to Athens for three hours.
b. The courier ran towards Athens for three hours.

This is because no matter how we decompose the event in question exhaustively into temporally shorter parts, at least one of them will either be discontinuous or fail to include the endpoint of the path of the event. Either way, this part will fail to qualify as *run all the way to Athens* (see Section 2.9). As for *run towards Athens*, no such problem arises because the endpoint of the path of an event in the denotation of this predicate does not have to be Athens.

5.5 Summary

This chapter has proposed a solution to the minimal-parts problem in terms of stratified reference. An atelic predicate *P* is defined as one such that every event in its denotation can be divided exhaustively into parts (“strata”) that are each in *P* and whose runtimes are properly included in the runtime described by the Key—the measure phrase of a *for*-adverbial. Since the Key differs across *for*-adverbials, the result is a variable notion of atelicity. An implementation in terms of algebraic closure was shown to be related to the concept of cover known from work by Schwarzschild (1996) and others on distributivity, and argued to be superior to previous formulations which use either universal or existential quantification over subintervals.

Aspect and space

6.1 Introduction

Most previous work on aspect, as well as this book so far, has focused on properties of temporal *for*-adverbials such as *for an hour* and *in an hour*, and on temporal properties of verbal predicates. This chapter starts from the observation that the telic/atelic opposition is not confined to the temporal domain, but is neatly replicated in space. Following Gawron (2004), I speak of temporal aspect and spatial aspect. Not all theories of aspect handle the extension from the temporal to the spatial domain equally well. This chapter shows that accounts of *for*-adverbials which are based on divisive reference, such as Krifka (1998), cannot be easily extended to the spatial domain. Alternative accounts based on the subinterval property, in the style of Dowty (1979), can be extended to the spatial domain without problems. Such accounts also include mine, since stratified reference can be seen as a generalization of the subinterval property.

Going from time to space takes aspect and telicity beyond their traditional usage. For example, Garey (1957), the source of the terms *telic* and *atelic*, defines a verbal predicate *V* to be atelic iff when one is interrupted while *V*ing, one can be said to have *V*ed.²³ The classical definition of aspect in Comrie (1976), “aspects are different ways of viewing the internal temporal constituency of a situation,” makes explicit reference to time. Likewise, the famous categorization by Vendler (1957) of verbal predicates into the four basic groups *state*, *activity*, *achievement*, *accomplishment* is based on time-related criteria such as interaction with temporal *for*- and *in*-adverbials and with the progressive.

As a consequence of extending the study of aspect to space, it becomes necessary to refine Vendler’s classical categorization. In modern implementations of this system, the difference between the four classes is partly determined by telicity: states and activities are atelic, achievements and accomplishments are telic (Verkuyl 1989).²⁴ Splitting

²³ This test is modeled on the distinction between *kinesis* and *energeia* in Aristotle’s *Metaphysics*, though the correspondence is not straightforward (Graham 1980).

²⁴ The other distinctions are as follows: states differ from activities in that they are taken to be “static,” and achievements differ from accomplishments in that they are taken to be punctual.

up telicity into spatial and temporal telicity means that we have to duplicate at least part of Vendler's categorization, and talk about spatial activities and accomplishments as Gawron (2005) does. The same is true for any of the other classification schemes that compete with Vendler's.²⁵ I do not dwell on this issue, because I regard Vendler classes as descriptive labels rather than analyses in their own right, in agreement with Verkuyl (1989, 2005) on this point. Vendler classes do not appear as primitives in this book, any more than they do in contemporary analyses of aspect such as Krifka (1998) and Verkuyl (1989, 1993).

While the traditional tests must be understood as diagnostics of temporal aspect, spatial aspect can be diagnosed by analogous means. Evidence for temporal and spatial aspect comes from a class of adverbial adjuncts which Moltmann (1991) calls *measure adverbials*. Moltmann applies this term to items like those in (1), taken from her paper.

- (1) a. John drank { wine / *a glass of wine } *for two hours*.
- b. Mary drew { pictures / *the picture } *until noon*.
- c. { Children / *1000 children } suffer from hunger *worldwide*.
- d. *Throughout the country* { women / *a hundred women } protested against the abortion law.

Measure adverbials in this sense include temporal as well as spatial *for*-adverbials. The examples in (2) are adapted from Gawron (2005).

- (2) a. The crack widens { for / *in } five meters.
- b. The crack widens two centimeters { in / *for } five meters.

Example (2a) has a stative interpretation on which it describes a crack whose width gradually increases over a stretch of five meters. By contrast, example (2b) cannot mean that the crack gradually widens by two centimeters over a stretch of five meters. It only has an irrelevant interpretation on which the crack's width over a stretch of five meters is two centimeters more than elsewhere, which arguably corresponds to result-oriented interpretations of *for*-adverbials such as *The Sheriff jailed Robin Hood for four years* (see Section 3.3).

The examples in (2) involve the degree achievement verb *widen*. In general, *widen* can be understood in two ways: either on a temporal interpretation, as involving reference to events in which an object such as an air balloon becomes wider and wider over time, or on an atemporal interpretation, as involving reference to events in which the width of an object increases gradually along a spatial axis. In this case, the object need not change its shape over time. To avoid confusion in connection with the word *event*, note that I also use this word in reference to states (see Section 2.4.3). Only the atemporal interpretation of *widen* and of other degree achievement verbs is relevant here. I have chosen the subject of the sentences in (2), *road*, to exclude any temporal

²⁵ Tatevosov (2002) has compiled a useful list of 16 such classifications.

interpretation that these sentences may have. Such an interpretation is implausible because roads do not normally change their shape over time.

Talmy (1996) analyzes examples such as *The scenery was rushing past me* in terms of “fictive motion.” While these examples are superficially similar to the sentences in (2), I agree with Gawron (2005) that fictive motion is not involved in this case. For example, the sentences in (2) cannot be placed into the progressive form, unlike Talmy’s example. In any case, implementing the concept of fictive motion in formal semantics seems more cumbersome to me than the straightforward extension of mereological approaches that I advocate here, and it would not capture the symmetrical behavior of temporal and spatial aspect.

Any attempt to view spatial *for*-adverbials as temporal *for*-adverbials in disguise will be challenged by examples that involve an area rather than a one-dimensional path:

- (3) The multiplicity of [a Chinese calligrapher’s] horizontal and vertical brush strokes, and their unending possibilities may be likened to sudden thunder and lightning which instantaneously flash for thousands of miles.²⁶
- (4) According to eyewitnesses, the ground was shaken for two miles around, and even the galleys tied up in the harbors felt the explosions through their wooden hulls.²⁷
- (5) [T]he police blocked streets for miles around [the museum].²⁸

Since a two-dimensional area cannot be mapped to a time interval in any obvious way, areas cannot easily serve as metaphoric descriptions of temporal intervals (Champollion 2015b). It seems more plausible to assume that either time or space can be encoded in a *for*-adverbial, and that not all *for*-adverbials are temporal.

Spatial aspect is not confined to degree achievement verbs, but is also exhibited by predicates like *meander* and *end* (Gawron 2005). As diagnosed by spatial adverbials, the former can be regarded as spatially atelic and the latter as spatially telic:

- (6) a. *The road meanders in a mile.
b. The road meanders for a mile.
- (7) a. The road ends in a mile.
b. *The road ends for a mile.

As noted by Gawron (2005), sentences like *The road meanders* and *The road ends* are stative, and applying the traditional, temporal-based notions of telicity to them

²⁶ Attested example, accessed May 26, 2016. http://www.ngansiumui.com/about/about_pub_ebook-callig_en.html.

²⁷ Attested example, accessed May 26, 2016. <http://www.historynet.com/the-guns-of-constantinople.htm>.

²⁸ Attested example, *New York Times*, *Iraq museum that was looted reopens, far from whole*, February 24, 2009. With thanks to Cleo Condoravdi (p.c.).

makes them temporally atelic, so the contrast between them cannot be described in terms of *temporal* telicity. However, what goes wrong in the spatial cases (1c), (1d), (2b), (6a), and (7b) does seem analogous to what goes wrong in the temporal cases (1a) and (1b). For example, (2a) does not just convey that the crack is wider at the end of the 5m stretch than at the beginning, but also that this widening is gradual and distributed over the length of the 5m stretch, in such a way that each part of the crack within that stretch widens. Similarly, it is plausible to link the status of (2b) to the fact that if the 5m stretch of the crack widens by 2cm in total, then the parts of the crack within that stretch do not each widen by 2cm. This is analogous to the explanation for (1a) in terms of aspectual composition, as found in Krifka (1998): the parts of an event of drinking wine are themselves events of drinking wine, but the parts of an event of drinking a glass of wine are not themselves events of drinking a glass of wine. This is so even though there are events which are both in the denotation of *drink wine* and in the denotation of *drink a glass of wine*. Again, there is an analogy with (2): among the states to which *The crack widens* applies, there are states to which *The crack widens two centimeters* applies.

Gawron views spatial aspect as involving change over a spatial rather than temporal axis. While I agree with the attempt to generalize boundedness from time to space, I will not adopt Gawron's treatment itself here. First, it would be difficult to integrate his analysis of aspect to the mereological framework which I have adopted. Second, Gawron attempts to characterize telicity in terms of cumulativity, but this is incompatible with the lexical cumulativity assumption which I need for independent purposes. On this assumption, all verbs have cumulative reference regardless of telicity (see Section 2.7.2).

Summing up, spatial measure adverbials provide us with a reason to dissociate the notion of aspect from time. Research into aspect must take seriously the analogous behavior of temporal and spatial *for*-adverbials, and any theory of aspect should account for both spatial and temporal *for*-adverbials in order to capture this parallel.

The rest of this chapter is organized as follows. Section 6.2 lays out the space of possibilities for a unified account of temporal and spatial (a)telicity: either we characterize the two notions by using a property that tests for application of the predicate in arbitrarily confined subregions of spacetime, such as cumulative or divisive reference, or we use a notion in which dimension (space or time) is a parameter and in which we test for application of the predicate in regions of spacetime that are constrained in only one dimension and can extend arbitrarily in the other dimensions. Based on the aspectual properties of predicates with one bounded and one unbounded argument, the section makes the case for the second option, which I call the strata-based approach. Section 6.3 presents Krifka (1998) as an example of the first option, which I call the subregion-based approach, and contrasts it with Dowty (1979) and my own account as two examples of the strata-based approach. Section 6.4 shows first intuitively and then in formal detail how only the second

option can account for the difference between spatial and temporal aspect. This book implements the second option because the higher-order property it uses, stratified reference, can be parametrized for time or space through its dimension parameter. Section 6.5 concludes.

6.2 Subregions versus strata

This section distinguishes two kinds of algebraic models of telicity and classifies them under the terms *subregion-based approach* and *strata-based approach*. These terms are chosen with reference to the metaphor introduced in Chapter 1. In this metaphor, events and other entities are thought of as occupying regions in an abstract space whose dimensions specify, among others, their spatial and temporal extent. Events can be both spatially and temporally extended at the same time (see Section 2.4.3). This opposes them to intervals, which correspond to one-dimensional lines (see Section 2.4.4). While the rest of this book considers thematic roles and measure functions to be among the dimensions of this abstract space, this part of the metaphor can be ignored here. Only the spatial dimensions and time matter for the present purpose, and the relevant subspace of the abstract space can therefore be thought of as spacetime. In fact, we will never have to consider more than one spatial dimension at a time, so the relevant subspace has only two dimensions and can be drawn on paper. I will make use of this fact when I provide illustrations.

By *strata* I mean parts of an event that are constrained in extent along one of these dimensions, and which extend up to the boundaries of the event along all the other dimensions. In a three-dimensional space, a stratum is formed like a layer, and the thinner it is, the closer it approximates a plane. In a two-dimensional space, the thinner a stratum is, the closer it comes to a line. By *subregions* I mean parts of an event that occupy arbitrary regions of spacetime, including regions that are constrained in all dimensions. For example, consider an event whose location extends fifty meters in one direction, and whose runtime is an interval of fifty minutes. A stratum of this event is either a part whose runtime is a proper subinterval of the fifty minutes but whose spatial extent is fifty meters, or a part whose spatial extent is a proper subinterval of the fifty meters but whose runtime is fifty minutes. A subregion is an arbitrary part of this event. Both the runtime and the spatial extent of a subregion could be proper parts of those of the whole event, but they do not have to be. It follows that any stratum of an event is also one of its subregions, but not vice versa.

I do not define strata and subregions formally because they are only used to build intuitions. They do not play a formal role in any of the theories I discuss, including my own, but the ideas they represent inform the following classification.

- I use the term **subregion-based approach** for implementations that characterize the difference between telicity and atelicity based on divisive reference or related

notions. Divisive reference holds of a predicate *P* if every part of any event in *P* is also in *P* (see Section 2.3.5). In our visual metaphor, divisive reference requires the event predicate to hold of every subregion of any event to which it applies. A faithful implementation of the subregion-based approach that is based on divisive reference would predict that it is impossible for one and the same predicate to be both spatially bounded and temporally unbounded, or vice versa, because an event predicate cannot be at the same time divisive and not divisive.

- The **strata-based approach** characterizes the difference between telicity and atelicity in a way that is parametrized on a dimension such as time or space. My own account of aspect implements the strata-based approach. The notion of stratified reference motivated in Chapters 4 and 5 abstracts over time and space via its dimension parameter. In our visual metaphor, stratified reference requires any event to which the event predicate applies to be subdivisible into strata that are constrained along the dimension specified by the dimension parameter, and run perpendicularly to it. The event predicate is then required to hold of every stratum of any event to which it applies. I will illustrate this in more detail in the following discussion. On the strata-based approach, the dimension parameter provides an additional degree of freedom compared with the subregion-based approach. This approach therefore makes it possible for predicates to be both spatially bounded and temporally unbounded or vice versa. For example, it is logically possible for one and the same predicate to have stratified reference with respect to runtime but not to location, or vice versa.

Each of the two characterizations of atelicity outlined has been implemented in various ways by different authors, in the shape of different lexical entries for *for*-adverbials. For example, the subregion-based approach is implemented in the work of Krifka (1998), while accounts that implement the strata-based approach include Dowty (1979) and Moltmann (1991) as well as this book. Here, I contrast prototypical examples of the two characterizations, namely Vendler (1957) and Bennett & Partee (1972). The next section contrasts the specific proposals by Krifka, Dowty, and myself. While neither Vendler nor Bennett & Partee proposed entries for *for*-adverbials, the simplicity of their proposals makes them useful to discuss for expository purposes before moving on to the more advanced implementations.

The essence of the subregion-based approach is present in Vendler's (1957) suggestion that (atelic) activities have the property that "any part of the process is of the same nature as a whole." Vendler contrasts this property with (telic) accomplishment predicates, which "proceed toward a terminus which is logically necessary to their being what they are." The property of atelic predicates that Vendler identifies corresponds to the model-theoretic notion of divisive reference. Although Vendler speaks

³⁰ Attested example, accessed May 25, 2016. <http://community.lawyers.com/forums/t/17235.aspx>.

While any of these examples raises similar problems for the subregion-based approach, I will focus on a minimal pair of sentences which are structurally similar to the ones just cited.³¹

- (9) a. John pushed carts all the way to the store for fifty minutes.
- b. John pushed carts all the way to the store for fifty meters.

Consider the following scenario:

- (10) BACK AND FORTH. John, an employee of a department store, is standing fifty meters away from the store in a parking lot. He has a few hundred shopping carts next to him, which he is supposed to push all the way to the store. Since these are too many carts to push all at once, he goes back and forth between the parking lot and the store, and each time he takes a few carts with him. After fifty minutes, his shift is over and he goes home.

In this situation, (9a) is true on an iterative interpretation of its verb phrase (as discussed in Section 2.7.2, I assume that iterativity comes for free as a consequence of lexical cumulativity). But (9b) is not, even though the distance John travels each time is fifty meters.³²

In Champollion (2010b), I assumed that (9b) is ungrammatical, and I treated (9a) and (9b) as differing in acceptability rather than in their meaning. However, Hans-Martin Gärtner (p.c.) offers a relevant example of a scenario in which both sentences may be judged acceptable and true (see Champollion 2015c):

- (11) ALONG THE WALL. The department store is one hundred meters wide. Along the store's wall there are carts every few meters that are at some distance from the wall. John walks along the store and on a stretch of fifty meters he pushes carts all the way to the wall.

In this scenario, (9b) is acceptable and true. If we also suppose that John works at a rate of just one meter per minute, so that he is done after fifty minutes, (9a) is acceptable and true as well.

The fact that temporal and spatial *for*-adverbials do not give rise to the same interpretations shows that there is no single property that can account for the semantics of all *for*-adverbials. If we want to characterize what happens in (9a) as well as

³¹ To prevent a partitive coercion of *to the store* into an unbounded predicate with a meaning similar to *towards the store*, I use the modifier *all the way*, which is a "true delimiter" in the sense of Smollett (2001).

³² A sentence like *John pushed carts all the way to the store* does not entail that John pushed more than one cart at a time; he could have pushed one cart at a time. This is consistent with the inclusive view on plurals, for which I have presented arguments in Section 2.6.2. In addition, the sentence entails that John pushed more than one cart in total. In the sentences in (9), the bare plural *carts* arguably acts as a dependent plural. That *for*-adverbials license dependent plurals can be seen from sentences like *John wore yellow neckties at night for a week*, which do not give rise to the conclusion that John ever wore more than one necktie at once.

(9b) in terms of (a)telicity, as was argued in the previous section, any approach to aspect needs to contain the equivalent of a parameter that can be set to space or time. The property of divisive reference does not contain such a parameter, so the subregion-based approach can be used to characterize such meaning differences only if it relies on a suitably amended version of divisive reference, or on another property altogether. The property of stratified reference contains such a parameter. I have called it the *dimension parameter*. Chapter 4 has motivated this parameter independently of aspect-related considerations.

Both the subregion-based and the strata-based approach presuppose that events and/or intervals have parts, so they can both be implemented within the subregion-based mereological framework presented in Chapter 2. Krifka (1989a, 1989b, 1992, 1998) and Kratzer (2007) implement the subregion-based approach in such a setting. Moreover, both theories can be formulated in an event-semantic setting. Dowty (1979), an early implementation of the strata-based approach, is formalized using neither event semantics nor mereology, but subsequent implementations are, including Hinrichs (1985) and Moltmann (1991).

6.3 Implementations of both approaches

This section presents concrete implementations of the subregion-based approach and the strata-based approach. The next section compares them with respect to how well they succeed in accounting for the different distribution of spatial and temporal measure adverbials. I present the proposal in Krifka (1998) as an example of the subregion-based approach. The strata-based approach is illustrated by Dowty (1979) and by my own proposal. For the purpose of this chapter, unlike in Chapter 5, the differences between Dowty's proposal (once it has been generalized to space) and my own are irrelevant, since both accounts are strata-based. Metaphorically speaking, these differences only concern the thickness of the strata: Dowty considers arbitrarily thin strata, while I relax that requirement.

6.3.1 Dowty (1979) and my proposal

Dowty (1979) essentially models *for*-adverbials as universal quantifiers ranging over the subintervals of some time interval. For example, a sentence like *John ran for an hour* is predicted to be true if John ran at each subinterval of a certain hour-long interval. I have reviewed Dowty's analysis of *for*-adverbials in Section 5.3.1. Since many of the details of his analysis are not relevant for the present purpose, I will not use Dowty's original entry, but a simplified version.

In Dowty's analysis, events and part structures are unnecessary, and Dowty himself does not make use of them. However, for purposes of comparison with Krifka's theory, it is useful to couch his analysis in a mereological subregion-based framework anyway

(see Moltmann 1991). To give all implementations a fair chance to model spatial aspect, I extend the background assumptions of the accounts I compare in this chapter in the following way: I assume the existence of a trace function τ that relates events to their runtime, and I assume that events are also mapped to spatial intervals by a location function σ that mirrors τ in all important respects. In particular, both σ and τ are assumed to be homomorphisms with respect to the sum operation. For example, the runtime of the sum of two events is the sum of their runtimes. For more details on these assumptions, see Section 2.5.2.

Casting both theories into the same framework is not only beneficial for comparison. Since I argue against the subregion-based approach, adopting the framework in which this theory was formulated makes sure that my criticism is targeted against the original theory and not against a potentially inaccurate translation. At the same time, my criticism of the subregion-based approach is not tied to the specific details of Krifka's implementation. It is targeted against the implicit premise of the subregion-based approach, and so it is likely to equally apply to other implementations of this approach.

The following is a simplified version of Dowty's entry that retains his basic insight, namely, that *for*-adverbials universally quantify over subintervals.

- (12) $\llbracket \text{for an hour} \rrbracket$ (Dowty, simplified)
 $= \lambda P \exists t [\text{hours}(t) = 1 \wedge \forall t' \subseteq t [AT(P, t')]]$

In Dowty's framework, the relation AT evaluates a proposition at an interval. In event semantics, AT can be reformulated in terms of τ . This definition is repeated here from Section 2.5.2.

- (13) **Definition: Holding at an interval**
 $AT(V, t) \stackrel{\text{def}}{=} \exists e [V(e) \wedge \tau(e) = t]$
 (An event predicate V holds at an interval t iff it holds of some event whose temporal trace is t .)

Given this definition, we can rewrite the translation in (12) as follows:

- (14) $\llbracket \text{for an hour} \rrbracket$ (Dowty, simplified)
 $= \lambda P \exists t [\text{hours}(t) = 1 \wedge \forall t' [t' \leq t \rightarrow \exists e [P(e) \wedge \tau(e) = t']]]$

I have replaced \subseteq by \leq because I assume a mereological rather than a set-theoretic structure for time (see Section 2.4.4).

After these modifications are carried out, Dowty's analysis produces representations like the following:

- (15) $\llbracket \text{John ran for an hour} \rrbracket$
 $= \exists t [\text{hours}(t) = 1 \wedge \forall t' [t' \leq t \rightarrow \exists e [\text{run}(e, \text{John}) \wedge \tau(e) = t']]]$
 (There is a time interval t which lasts an hour and John runs at each of its parts.)

- (16) $\llbracket \text{John ran for a mile} \rrbracket$
 $= \exists t [\text{miles}(t) = 1 \wedge \forall t' [t' \leq t \rightarrow \exists e [\text{run}(e, \text{John}) \wedge \sigma(e) = t']]]$
 (There is a spatial interval t which spans a mile and John runs at each of its parts.)

Dowty's analysis can be classified as an implementation of the strata-based approach because he lets *for*-adverbials introduce a test that applies to subintervals. This feature is also implemented in my own proposal, as presented in Chapters 4 and 5. My translation for the temporal adverbial *for an hour* is shown in (17). As a reminder, $\gamma(\tau, e)$ was defined in Section 4.5.4 as the predicate which holds of intervals that are properly included in the runtime of e .

- (17) $\llbracket \text{for an hour} \rrbracket$ (my proposal, repeated from Section 4.7)
 $= \lambda P_{\langle v, t \rangle} \lambda e : \text{SR}_{\tau, \gamma(\tau, e)}(P)(e).$
 $P(e) \wedge \text{hours}(\tau(e)) = 1 \wedge \text{regular}(\tau(e))$

This entry results in logical representations like the one in (18).

- (18) $\llbracket \text{John ran for an hour} \rrbracket$
 $\exists e : \text{SR}_{\tau, \gamma(\tau, e)}(\llbracket \text{John run} \rrbracket)(e).$
 $\llbracket \llbracket \text{John run} \rrbracket(e) \wedge \text{hours}(\tau(e)) = 1 \wedge \text{regular}(\tau(e))$
 (There is a regular running event e whose runtime measures an hour, and it is presupposed that *John run* stratifies e with dimension τ and granularity $\gamma(\tau, e)$.)

Given the definition of stratified reference in Section 4.6, the presupposition of (18) can be expanded as in (19).

- (19) $\text{SR}_{\tau, \gamma(\tau, e)}(\llbracket \text{John run} \rrbracket)(e) \Leftrightarrow$
 $e \in * \lambda e' \left(\begin{array}{l} \tau(e') < \tau(e) \wedge \\ \llbracket \llbracket \text{John run} \rrbracket(e') \end{array} \right)$
 (The hour-long event e can be divided into parts each of which is in $\llbracket \text{John run} \rrbracket$ and has a runtime properly included in the runtime of e .)

The corresponding spatial translations are exactly parallel. Spatial *for*-adverbials impose a stratified-reference requirement that is parametrized for σ (spatial extent) instead of τ (runtime) and is otherwise equivalent. Other spatial measure adverbials, such as *worldwide* in (1c) and *throughout the country* in (1d), can be represented in similar ways.

- (20) $\llbracket \text{for a mile} \rrbracket$ (my proposal)
 $= \lambda P_{\langle v, t \rangle} \lambda e : \text{SR}_{\sigma, \gamma(\sigma, e)}(P)(e).$
 $P(e) \wedge \text{miles}(\sigma(e)) = 1 \wedge \text{regular}(\sigma(e))$
- (21) $\llbracket \text{John ran for a mile} \rrbracket$
 $\exists e : \text{SR}_{\tau, \gamma(\sigma, e)}(\llbracket \text{John run} \rrbracket)(e).$

$[\llbracket \text{John run} \rrbracket(e) \wedge \text{miles}(\sigma(e)) = 1 \wedge \text{regular}(\sigma(e))]$

(There is a regular running event e whose spatial extent measures a mile, and it is presupposed that *John run* stratifies e with dimension σ and granularity $\gamma(\sigma, e)$.)

(22) $\text{SR}_{\sigma, \gamma(\sigma, e)}(\llbracket \text{John run} \rrbracket)(e) \Leftrightarrow$

$e \in {}^*\lambda e' \left(\begin{array}{l} \sigma(e') < \sigma(e) \wedge \\ \llbracket \text{John run} \rrbracket(e') \end{array} \right)$

(The mile-long event e can be divided into parts each of which is in $\llbracket \text{John run} \rrbracket$ and has a spatial extent properly included in the spatial extent of e .)

6.3.2 Krifka (1998)

Krifka (1998) translates *for*-adverbials using a parametrized version of divisive reference, and therefore instantiates the subregion-based approach. Krifka's position has evolved over the years: In Krifka (1989b: ch. 3), *for*-adverbials impose the requirement that predicates be divisive and strictly cumulative (i.e. they are cumulative and they have at least two distinct entities in their denotation), and the predicate that is output by the *for*-adverbial must not be cumulative. In Krifka (1989a), atelic predicates are considered to be "strictly cumulative or at least non-quantized." Krifka (1998) contains the most recent entry for *for*-adverbials published by the author.

The translation given in Krifka (1998) for *for an hour* and its use in *John ran for an hour* are shown in (23) and (24). I present Krifka's entry here in its original form to show that the reason it fails is not due to a simplification of mine. I explain Krifka's notation below.

(23) $\llbracket \text{for an hour} \rrbracket$ (Krifka)

$= \lambda R. \lambda x. \lambda e. [R(x, e) \wedge H'(e) = 1 \wedge$
 $\partial \exists e' [e' <_{\tau} e \wedge \forall e'' [e'' \leq_{\tau} e \rightarrow R(x, e'')]]]$

(24) $\llbracket \text{John ran for an hour} \rrbracket$ (Krifka)

$= \exists e. [\text{run}(\text{John}, e) \wedge H'(e) = 1 \wedge$
 $\partial \exists e' [e' <_{\tau} e \wedge \forall e'' [e'' \leq_{\tau} e \rightarrow \text{run}(\text{John}, e'')]]]$

The translation in (23) consists of an assertion and of a presupposition, which is indicated by the unary operator ∂ (Beaver 2001). The translation takes a predicate of individuals and events R , an individual x , and an event e . In (24), the predicate R is assumed to have been contributed by the verb phrase, x by the subject, and e by existential closure. The entry states that the number of hours of the runtime of e is one. This is expressed by the function H' that maps entities to their runtimes in hours (with some special provisions for noncontinuous events that are irrelevant for the present point).

The presupposition of (23) makes use of a parametrized part-of relation \leq_{τ} . Krifka defines that $e' \leq_{\tau} e$, read " e' is a temporal subevent of e ," holds iff $e' \leq e$ and there

is another part e'' of e whose runtime $\tau(e'')$ does not overlap with $\tau(e')$.³³ This is the case if the runtime of e' is a proper part of the runtime of e . Thus, Krifka relativizes the part-of relation to time. Note that nothing in the definition requires e and e' to coincide in their spatial extent or in any other way. In particular, e' could also have a smaller spatial extent than e .

The presupposition of (23) states that the predicate R denoted by the verb phrase must relate the subject to all the temporal subevents of the event e . I will refer to this part of the presupposition as the *divisiveness clause*, because it essentially requires the predicate to have divisive reference, except that \leq is replaced by \leq_τ . The presupposition also requires that the event has temporally shorter subevents. I call this the *existence clause*. The purpose of this clause is to exclude quantized telic predicates like *eat an apple*. These predicates vacuously satisfy the divisiveness clause, because an event to which *eat an apple* applies arguably has no subevents—and, in particular, no temporally shorter subevents—that would also be in the denotation of *eat an apple*.

Although Krifka does not deal with spatial *for*-adverbials, his theory can be straightforwardly extended to them by stating that spatial *for*-adverbials check that each spatial (as opposed to temporal) subevent of the sum event is in the denotation of the main clause predicate. This is shown in (25), which uses a relation \leq_σ or “is a spatial subevent of,” to be understood in a parallel way to \leq_τ .

- (25) $\llbracket \text{for a mile} \rrbracket$ (based on Krifka)
 $= \lambda R. \lambda x. \lambda e. [R(x, e) \wedge M'(e) = 1 \wedge$
 $\partial \exists e' [e' <_\sigma e \wedge \forall e'' [e'' \leq_\sigma e \rightarrow R(x, e'')]]]$

6.4 Comparing the approaches

In this section, I describe a case for which the proposal in Krifka (1998) makes the wrong predictions, and which the proposal in Dowty (1979) and my own proposal handle correctly. Consider again sentence (9a), repeated here:

- (26) John pushed carts all the way to the store for fifty minutes.

The events of which this sentence holds are sums of events in which John goes back and forth many times. For the present purpose, I assume that the predicate *John push carts all the way to the store* is interpreted as in (27) (see also Section 2.9 for my assumptions concerning the semantics of *to the store*; I assume that *all the way* is redundant here). Krifka (1998) assumes a similar representation, although I abstract away from some irrelevant details here.

³³ Krifka actually writes $\leq_{H'}$ rather than \leq_τ . However, the function H' is not used in his definition of \leq_τ . I write Krifka's relation as \leq_τ because I find this notation more intuitive when it is used in parallel with its spatial counterpart, which I write \leq_σ .

- (27) $\lambda e[*push(e) \wedge *agent(e) = j \wedge end(\sigma(e)) = the.store]$
 (True of any potentially plural pushing event whose agent is John and whose spatial extent is an interval whose end is the store.)

I will first illustrate the different behavior of the two approaches intuitively with figures, and then discuss more formally what goes wrong with Krifka's account. I will overlay the constraints that the two approaches impose onto a schematic representation of a sum event of which (26) holds. More specifically, I will assume from now on that the sum event verifies BACK AND FORTH, where John starts out fifty meters away from the store with all the shopping carts next to him. This schematic representation is shown in Figure 6.1. It uses the visual metaphor described in Chapter 1 and Section 6.2. This figure is a spacetime diagram with time on the vertical axis, and with the path or spatial interval to which *all the way to the store* refers along the horizontal axis. I have taken certain liberties here for expository purposes. Strictly speaking, in a spacetime diagram, a stationary object like the store should be represented as a vertical line, and John and his carts should be represented as zigzagging spacetime "worms." I have used more recognizable representations. One can imagine Figure 6.1 as taken from a vertical filmstrip that has been recorded from a perpendicular angle to John's path. The men and carts can be thought of as individual frames, each of which represents an event in which John pushes carts all the way to the store.

Figure 6.2 contrasts two ways of breaking down this sum event into subevents, corresponding to the two theories. Figure 6.2a illustrates the strata-based approach: the event is divided into subevents along the temporal axis. Each of these events is tested for whether it qualifies, that is, whether it is in the denotation of the predicate *John push carts all the way to the store*. A checkmark indicates that an event qualifies, a cross shows that it does not. A picture in which there are only checkmarks and no crosses represents a case in which the theory predicts the sentence to be acceptable. Since this is the case in Figure 6.2a, this is a point for the strata-based approach.

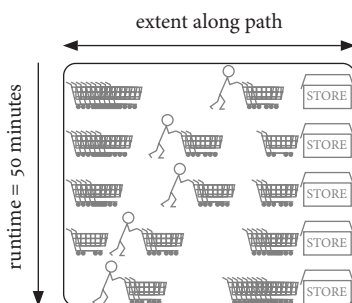


FIG. 6.1 An event in the denotation of (26).

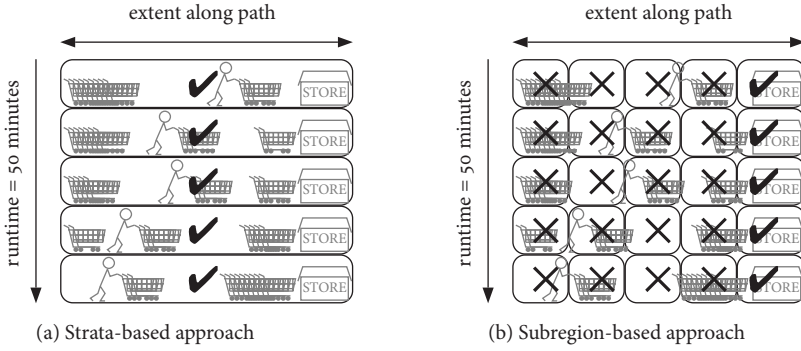


FIG. 6.2 Comparison of the strata-based and subregion-based approaches.

The subregion-based approach is illustrated in Figure 6.2b. As shown by the abundance of crosses in this figure, this approach breaks down the sum event too much by insisting on considering all subevents (or at least all temporal subevents, even those whose spatial extent is smaller). Most of these subevents have locations that are remote from the location of the store, and are therefore not in the denotation of the predicate *John push carts all the way to the store*. The restriction to temporally shorter subevents does not prevent this. The subregion-based approach in general, and Krifka's theory in particular, wrongly rules out sentence (26).

6.4.1 The subregion-based approach

Following this intuitive presentation, let me now show in detail how the account in Krifka (1998) wrongly rules out sentence (26) even though it is more sophisticated than a naïve implementation of the subregion-based approach. I will argue that the presupposition predicted by Krifka's entry in (23) is violated by the predicate denoted by (26). As a result, Krifka's account wrongly predicts that (26) should be unacceptable due to presupposition failure. His prediction is that (26) presupposes that any event e in its denotation satisfies the following condition:

$$(28) \quad \exists e' [e' <_{\tau} e] \wedge \forall e'' [e'' \leq_{\tau} e \rightarrow e'' \in \llbracket \text{John push carts all the way to the store} \rrbracket]$$

As a reminder, I call the first conjunct of this presupposition the *existence clause* and the second conjunct the *divisiveness clause*. To argue that this presupposition is violated, I will proceed in four steps. First, I will argue that sentence (26) applies to a proper sum event, i.e. an event that has subevents. This means that the existence clause of the presupposition is satisfied, which means that the problem must lie in the divisiveness clause and ultimately in Krifka's reliance on divisive reference. Second, I will argue that some of these shorter subevents are not in the denotation of the

predicate *John push carts all the way to the store*. Third, I will argue that some of these subevents stand in the relation \leq_τ and therefore \leq to the sum event (they are parts of the sum event and their runtime is shorter than that of the sum event). Finally, I will argue that these subevents cannot be excluded on pragmatic grounds from the range of the universal quantifier in (28). These four steps together have the consequence that the presupposition in (28) is not satisfied.

I start by arguing that (26) applies to a sum event that has subevents. I have assumed that a verbal predicate that has been modified by *all the way to the store* only applies to events whose spatial extent ends at the store (see Section 2.9). This assumption is intuitively plausible, and corresponds to the account of spatial prepositional phrases in Krifka (1998). Suppose that (26) is uttered in a scenario in which John repeatedly pushed some carts from location A to the store. (In a pragmatically plausible scenario, he pushes a different set of carts on each of his back-and-forth trips, but nothing depends on this.) Call B the point halfway between location A and the store. Standard assumptions of algebraic event semantics (see Section 2.7) commit us to the (intuitively plausible) claim that (26) entails the existence of a sum event e which can be divided at least into two complex subevents: an event e_1 in which John pushed carts (iteratively) from A to B and an event e_2 in which he pushed carts (iteratively) from B to the store. I will leave e_2 aside and concentrate on the event e_1 . This event does not itself qualify as an event of John pushing carts all the way to the store, because its spatial extent does not include the store. To see that e entails the existence of e_1 , observe that (26) entails (29), and moreover (29) entails (30):

(29) John pushed carts all the way to the store.

(30) John pushed carts halfway towards the store.

These sentences are of course less informative than (26) because they describe parts of the scenario evoked by it, so it is not immediately obvious that the entailment relations that I have mentioned hold. For example, one might take (30) to be false in such a scenario because it conveys that the carts in question did not reach the store. However, that the carts did not reach the store is not a part of the literal meaning, but an implicature. This is clear in contexts such as questions, where implicatures are usually not computed. For example, when we turn (30) into a question, it is possible to answer it affirmatively with (29), but the converse is not possible:

(31) a. Did John push carts halfway towards the store?
b. { Yes / *No }—in fact, he pushed carts all the way to the store.

(32) a. Did John push carts all the way to the store?
b. { No / *Yes }, he pushed carts halfway towards the store.

I conclude that the entailment relations do indeed hold. Each of these sentences entails the existence of an event, and we can plausibly model the fact that (26)

entails (30) by assuming that any event denoted by (26) has an event denoted by (30) as one of its parts. This means that (26) indeed denotes a sum event, as I have claimed.

The second step is to argue that the predicate denoted by (29), which applies to the sum event of (26), call it e , does not apply to the sum event of (30), call it e_1 . This is easy to see, given that we know that (29) entails (30) but not vice versa. If *John pushed carts all the way to the store* applied to the event in (30), then (30) should entail (29). The intuition behind this reasoning is that an event of pushing carts halfway towards the store neither is nor entails an event of pushing carts all the way to the store.

Third, I argue that $e_1 \leq_\tau e$, that is, e_1 is a part of e and there is another part of e whose runtime does not overlap with the runtime of e_1 . This is how Krifka expresses that e_1 has a shorter runtime than e . As mentioned, that $e_1 \leq e$ is intuitively plausible. Any assumption to the contrary would make it very difficult to explain why (26) entails (30). Moreover, observe that e_1 is a proper part of e , i.e. it is not identical to e . This is so because (30) does not entail (29). By the Unique Separation axiom (see Section 2.3.1), there is an event e_2 that is a proper part of e and does not overlap with e_1 , so that the sum of e_1 and e_2 is e . Intuitively, this models the following fact: e is a sum event in which John pushed carts from A to the store; e_1 is a part of e in which John pushed carts from A halfway towards the store, that is, from A to B; and e_2 is an event in which John pushed carts from B to the store, that is, the result of subtracting e_1 from e . Now, given the fact that John cannot be in two different locations at the same time, the runtimes of e_1 and e_2 do not overlap. This entails $e_1 \leq_\tau e$, which I set out to show.³⁴

Summing up, we have the following situation. A *for*-adverbial can modify a sentence, namely (26), whose sum event e has a proper part e_1 that is not in the denotation of the sentence. This is in contradiction to divisive reference. Moreover, e_1 has a shorter runtime than the sum event e , therefore not only $e_1 < e$ holds but also $e_1 <_\tau e$. This is in contradiction to the divisiveness clause, by which Krifka (1998) implements the subregion-based approach, as we have seen in (23). I will refer to e_1 as the “offending event” since it violates the divisiveness clause. We will see later that the strata-based approach fares better because e_1 does not lead to a violation of stratified reference.

Finally, let us consider what additional assumptions would be necessary to rescue the subregion-based approach. Although this is not always explicitly stated, universal quantifiers in semantic representations are sometimes assumed to be restricted to “relevant” entities. The following classical example, taken from Kratzer (1989), illustrates the point. We have an orchard whose trees are all laden with wonderful apples. A man who wants to buy the orchard asks us whether all its trees are apple trees. We answer:

³⁴ The runtimes of each of these events will be discontinuous in scenarios like BACK AND FORTH where John goes back and forth between A and the store and pushes the carts all the way to the store one by one or at least little by little. This does not affect the main point.

“Yes, and every tree is laden with wonderful apples.” It is clear from context that the universal quantifier supplied by *every tree* is implicitly restricted to the trees in the orchard. Otherwise, the sentence would be false because there are many trees in the world which are not apple trees and cannot have any apples. The relevance issue is a bit of a wildcard, because no commonly accepted semantic theory provides clear criteria for deciding whether a given entity is relevant. However, a common intuition is that all irrelevant entities must be “outside the situation” in which the sentence is understood. In the orchard example, most if not all of the trees inside the orchard are relevant from an intuitive point of view since the truth of the sentence depends on any one of them having apples, while the trees outside the orchard are irrelevant. This intuition is formally implemented in situation semantics (Kratzer 1989).

In our event-based framework, the situation in which a sentence is understood can be thought of as the sum event over which a sentence existentially quantifies (Dekker 1997, Kratzer 2016). According to this view, the relevant events in the example discussed above are the subevents of the sum event *e*. Event *e*₁, the “offending event” whose location does not contain the location of the store, is indeed a subevent of *e* and is therefore relevant. As shown in Figure 6.2b, there are many other such offending subevents. It would be difficult to explain why none of these subevents should be seen as relevant.

However, let us grant that an argument to the contrary can be made, and that the only relevant events turn out to be those subevents of *e* whose location contains the store. In Figure 6.2b, these subevents are the ones whose right-hand boundary is also the right-hand boundary of the sum event. These subevents all have checkmarks. This indicates that they all qualify as *push carts all the way to the store*, which is the case because their location contains the store. It would then follow that all relevant subevents of *e* would qualify as *push carts all the way to the store*, and the presupposition of (26) would indeed be satisfied. This would appear to rescue the subregion approach.

I see three problems with this line of thinking. First, there is no reason to assume that these subevents are indeed irrelevant for the felicity and truth conditions of the sentence, and that so many of them should be. If they were, we should expect most trees inside our orchard to be irrelevant for the truth conditions of *Every tree is laden with apples* as well. Second, there is no explanation why (26) requires an iterative interpretation, as in BACK AND FORTH. Suppose John starts pushing a set of carts towards the store all at once, and stops doing anything once he has arrived there. Although the locations of most subevents of such an event do not contain the store, those whose location does contain the store all qualify as events in which John pushes carts all the way to the store. If these are the only relevant subevents, the sentence should be modifiable by any temporal *for*-adverbial. Finally, if we can appeal to contextual restriction to remove offending events as needed, there is no clear explanation why spatial *for*-adverbials make (26) unacceptable in BACK AND

FORTH but acceptable in ALONG THE WALL, while temporal *for*-adverbials show the opposite pattern. I conclude that an appeal to domain restriction is unlikely to save the subregion-based approach.

6.4.2 The strata-based approach

I will now argue that the strata-based approach provides a handle on the observations presented in Section 6.2: that the sentence *John pushed carts all the way to the store* is compatible with a temporal *for*-adverbial; that this adverbial enforces an iterative interpretation, as in BACK AND FORTH; and that a spatial *for*-adverbial is incompatible with BACK AND FORTH but compatible with ALONG THE WALL. The strata-based approach was illustrated in Figure 6.2a. The facts just mentioned are illustrated by example (9), repeated here:

- (33) a. John pushed carts all the way to the store for fifty minutes.
b. John pushed carts all the way to the store for fifty meters.

I start with the first observation: the sentence is compatible with a temporal *for*-adverbial. Take any event in the denotation of (33a) and call it e . As implemented in Dowty's proposal, the strata-based approach checks if every subinterval of the runtime of e is the runtime of an event in the denotation of *John push carts all the way to the store*. This condition is of course problematic, because there are no infinitely short events that qualify as pushing events, let alone events of pushing carts all the way to the store. However, Chapter 5 shows that the strata-based approach can be maintained in spite of this minimal-parts problem. My implementation of the strata-based approach checks if there is a way to divide up the sum event in (33a) along the temporal axis into strata that are all in the denotation of the main clause predicate. The maximum duration of these strata depends on the length of the temporal interval associated with the measure phrase of the *for*-adverbial. In the case of (33a), their runtime will have to be properly contained in the fifty-minute runtime of the whole event—and typically they will in fact be much shorter, say, five minutes.

- (34) $SR_{\tau, \gamma(\tau, e)}(\llbracket \text{John push carts all the way to the store} \rrbracket)(e) \Leftrightarrow$
 $e \in * \lambda e' \left(\begin{array}{l} \tau(e') < \tau(e) \wedge \\ \llbracket \text{John push carts all the way to the store} \rrbracket(e') \end{array} \right)$
 (The fifty-minute event e can be divided into parts each of which is in $\llbracket \text{John push carts all the way to the store} \rrbracket$ and has a runtime properly included in the runtime of e .)

Because the strata-based approach does not quantify over all temporally shorter subevents, it places a weaker requirement on the main-clause predicate than Krifka's account does. We have seen that the sum event of (33a), e , can be divided into two parts e_1 and e_2 such that e_1 does not qualify as an event in which John pushed carts all the way to the store. More specifically, e_1 represented the sum of all subevents in which

John pushed carts halfway towards the store. In other words, the existence of the parts e_1 and e_2 represents a way of dividing up e along the spatial axis. The existence of e_1 , the “offending event,” was problematic for Krifka (1998), but it is without consequence for my proposal. All that needs to be checked is whether there exists a way of dividing e into subevents along the temporal axis, such that these subevents are each in the denotation of the main clause predicate. Once it has been established that such a division exists, it does not matter whether or not there are also other ways of dividing e into parts to which the main clause predicate does not apply.

To check whether e can be divided up in the specified way, we need to check the scenarios with which (29) is compatible. Example (29) by itself does not specify whether John pushed carts all the way to the store little by little (iteratively, that is, with many trips back and forth as in *BACK AND FORTH*) or all at once; but as mentioned, adding a temporal *for*-adverbial as in (33a) forces an iterative interpretation. On my proposal, this contrast is expected. A scenario in which John pushed the carts all the way to the store little by little entails that the main-clause predicate in (29) is true within many consecutive time intervals—whatever time it takes John to push a set of carts all the way to the store and to go back to the origin. A scenario in which John pushed the carts to the store all at once would not be compatible with such an entailment. This entailment is exactly the one that is also required by *for*-adverbials on the strata-based approach.

The strata-based approach not only accounts for the acceptability and licensed entailments of (33a). It also explains the difference in entailments between the temporal *for*-adverbial in (33a) and its spatial counterpart in (33b). The reason for the difference between these two sentences is easily explained on the strata-based approach, given any background theory that makes the uncontroversial prediction that the PP *all the way to the store* is only true of events whose spatial extent ends at the store. The idea is this: While the temporal *for*-adverbial in (33a) tests whether the main-clause predicate is true within temporal subintervals that are properly included in the fifty minutes in question, the spatial *for*-adverbial in (33b) tests whether the main-clause predicate is true within proper spatial subintervals of the fifty meters in question. In *BACK AND FORTH*, some of these subintervals will not spatially contain the location of the store, i.e. they will not contain the endpoint of the spatial interval at which the sum event is located (see Figure 6.3). Any event that is located in such a subinterval will not qualify as *push carts all the way to the store*, i.e. it will not be in the denotation of the main-clause predicate. This means that there is no way of dividing up a sum event denoted by *John pushed carts all the way to the store* that fits *BACK AND FORTH* along the dimension of the spatial *for*-adverbial in (33b) that would conform to the requirements of the *for*-adverbial. In *ALONG THE WALL*, it is the opposite.

In general, the strata-based approach can easily capture the requirements of spatial *for*-adverbials we have seen in examples like (6b) and (7b), repeated here:

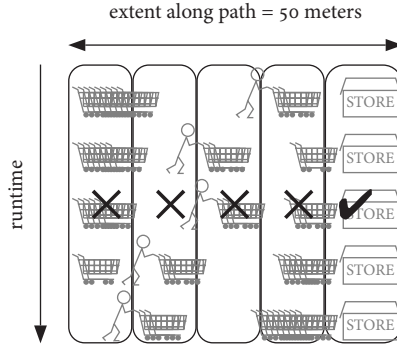


FIG. 6.3 Analyzing *John pushed carts all the way to the store for fifty meters* on the strata-based approach.

- (35) a. The road meanders for a mile.
 b. *The road ends for a mile.

Examples like (35a) do not require the road in question to meander throughout its entire length, just as the sentence *John walked for an hour* does not require John to walk throughout his entire lifetime. To represent this fact, I assume that the relevant sentences involve underlying states and that these states may be spatiotemporally extended, just like events in the narrow sense (see Section 2.4.3). The state that verifies *The road meanders* will have the same spatial extent as the meandering part of the road. Using the translation of *for a mile* in (20), we obtain the following definedness condition for (35a), where e is the state that verifies the sentence:

$$(36) \quad \text{SR}_{\tau, \gamma(\sigma, e)}(*\text{meander})(e) \Leftrightarrow e \in * \lambda e' \left(\begin{array}{l} * \text{meander}(e') \wedge \\ \sigma(e') < \sigma(e) \end{array} \right)$$

(The meandering state e in question can be exhaustively divided into parts (strata) which are also meandering states and whose spatial extents are properly included in the spatial extent of e .)

6.5 Summary

Temporal and spatial *for*-adverbials impose analogous constraints on the predicates with which they combine. Any theory of telicity should therefore account for temporal as well as spatial *for*-adverbials. Since the two kinds of adverbials differ in the scenarios they are compatible with, any theory of telicity should be able to classify a given predicate as temporally telic but spatially atelic or vice versa. This situation supports a parametrized notion of the telic/atelic opposition, where the parameter is set either to time or to a spatial dimension.

The parametrized nature of aspect is expected within the general picture of this book. As argued in Chapter 4, the fact that a predicate can be distributive on its agent “dimension” without being distributive on its theme “dimension” or vice versa supports a parametrized view of distributivity. In connection with measurement, I have argued in Chapter 4 that the singular/plural and count/mass distinctions that are operational in the pseudopartitive construction need to be parametrized to a “dimension” like width or diameter. The concept of stratified reference subsumes this picture. The dimension parameter of stratified reference gives us a handle on the distinction between spatial and temporal aspect.

I have argued that subinterval-property-based or “strata-based” approaches to atelicity, modified as in Chapter 5, generalize to the spatial case in a better way than approaches which are based on divisive reference. On strata-based approaches, an atelic predicate is one that holds of subevents which are constrained in size along the dimension with respect to which the predicate is atelic, but not along other dimensions. Approaches based on divisive reference impose conditions which are too strong, because they also require the predicate to hold of subevents whose size is small along all dimensions. Once again, the picture is subsumed by the concept of stratified reference. According to this concept as applied in Chapter 4, a predicate that is distributive on its agent “dimension” is not required to apply to subevents that are constrained on their theme dimensions—that is, it is not required for an agent-distributive predicate to be distributive on its theme as well.

Measure functions

7.1 Introduction

In Chapters 4, 5, and 6, we have seen how stratified reference characterizes the telic/atelic opposition and how it allows us to avoid the minimal-parts problem and generalize to spatial aspect. In this chapter, I turn to measurement in natural language. I consider the formal properties of measure functions and the way in which distributive constructions constrain these properties. Several measurement-related constructions turn out to be sensitive to a number of oppositions that can all be characterized by using stratified reference. Aside from measure adverbials like *for an hour*, I focus on pseudopartitives (Selkirk 1977).

As in Chapters 4 and 6, I draw attention to circumstances involving several dimensions. In this case, the dimensions correspond to different measure functions. I show that the constraint on measure functions only applies along one dimension at a time. I suggest that this behavior has the same source as the atelicity constraint in *for*-adverbials and the distributivity constraint in adverbial-*each* constructions.

My background assumptions on pseudopartitives and measure functions were laid out in Sections 2.5.3 and 3.2. Section 2.5.2 also introduced trace functions, and Section 3.2 assumed that event pseudopartitives involve trace functions. Measure functions relate individuals and events to their degrees on various scales, while trace functions relate events to their runtimes and locations. The rest of this book distinguishes them from measure functions mainly for technical reasons, for example because they have different types. Trace functions behave analogously to measure functions with respect to the facts I will discuss here. Therefore, this chapter does not distinguish between measure functions and trace functions.

Section 7.2 reviews constraints that pseudopartitives impose on the measure functions and on the substance nouns that occur in them, focusing on the distinction between extensive and intensive measure functions, and argues that these constraints are also operative in *for*-adverbials. Section 7.3 discusses previous accounts of the constraints on measure functions in pseudopartitives, focusing on Krifka (1998) and Schwarzschild (2006). Section 7.4 shows that the insights behind these accounts can be subsumed by stratified reference and the Distributivity Constraint. Section 7.5 concludes and draws connections to previous chapters.

7.2 Constraints on measure functions

As the examples in (1) illustrate, pseudopartitives can be used to talk about substances and events in terms of their measurement along various dimensions:

- | | | |
|-----|----------------------------------|-----------------------|
| (1) | a. five pounds of rice | <i>weight</i> |
| | b. five liters of water | <i>volume</i> |
| | c. five hours of talks | <i>temporal trace</i> |
| | d. five miles of railroad tracks | <i>spatial extent</i> |

Chapter 4 has suggested that the Distributivity Constraint applies to all distributive constructions. This leads to the prediction that pseudopartitives and other constructions, such as *for*-adverbials, should behave in similar ways.

There are at least two semantic parallels between pseudopartitives and *for*-adverbials. First, both reject predicates that fail to apply to the parts of the entities and events in their denotation. This category includes telic predicates in the case of *for*-adverbials, as we have seen in the previous chapters, and count nouns in the case of pseudopartitives, as illustrated in (2). It is not possible to use a pseudopartitive like the one in (2c) to describe a single book whose weight is five pounds.

- | | | |
|-----|-------------------------|---|
| (2) | a. five pounds of books | |
| | b. five pounds of rice | |
| | c. *five pounds of book | <i>(unacceptable with "book" as a count noun)</i> |

Several irrelevant interpretations may be available for (2c) when the identity of the books in question is not at issue or not recoverable. For example, it may be interpreted as describing five pounds of books in a situation where the books are sold by weight, or five pounds of pulp that results from shredding books. In such cases I assume that the count noun *book* has been coerced to a mass noun (see Section 2.6). I leave these interpretations aside since they correspond to (2b), where a mass noun is used.

The second semantic parallel between pseudopartitives and *for*-adverbials is that both of them reject measure functions whose value tends to stay constant across the parts of any object or event they measure. Examples of such functions are speed, as illustrated by example (3b), and temperature, as in example (4b).

- | | |
|-----|---------------------------------------|
| (3) | a. five hours of running |
| | b. *five miles an hour of running |
| (4) | a. thirty liters of water |
| | b. *thirty degrees (Celsius) of water |

Although I will focus on pseudopartitives in this chapter, several other constructions behave analogously. For example, when comparative determiners are used with substance mass nouns, they are underspecified as to what measure function is involved

(Schwarzschild 2006). Thus, *more rope* can mean “a longer portion of rope” or “a heavier portion of rope” in different contexts, depending on what is relevant. It is not possible, however, to use comparative determiners to compare two amounts of rope by temperature. In other words, it is not possible to come up with a context in which *more rope* can be used felicitously to mean “a warmer portion of rope.” True partitives are another construction that behaves similarly to pseudopartitives in terms of rejecting certain measure functions, as shown here:

- (5) a. *five degrees Celsius of the water in this bottle *temperature
 b. *five miles per hour of my driving *speed

From the examples so far, one might think that certain measure functions are never acceptable in pseudopartitives. But in fact, measure functions that are usually unacceptable can be made acceptable when the substance noun is chosen in the right way, as in the following attested example:

- (6) The scientists from Princeton and Harvard universities say just two degrees Celsius of global warming, which is widely expected to occur in coming decades, could be enough to inundate the planet.³⁵

This means we cannot simply categorize measure functions as acceptable or unacceptable per se. What matters is whether they are acceptable on the set denoted by the substance noun of the pseudopartitive in which they appear.

7.3 Previous accounts

The fact that pseudopartitives accept certain measure functions but reject others has previously been linked to the measure-theoretic properties of these measure functions. Krifka (1989a, 1998) and Schwarzschild (2002, 2006) note that the constraint corresponds to a distinction commonly made in measurement theory and in physics, namely the one between extensive and intensive measure functions. In physics, an extensive measure function is one whose magnitude is additive for subsystems; an intensive measure function is one whose magnitude is independent of the extent of the system (Krantz, Luce, Suppes & Tversky 1971; Mills, Cvitaš, Homann, Kallay & Kuchitsu 2007). For example, when one considers the system consisting of the water in a tank, volume is an extensive measure function because it is additive, meaning that the volume of the water as a whole is greater than the volume of any of its proper parts. But temperature is intensive with respect to this system because the temperature of the water as a whole is no different from the temperature of its proper parts. Krifka (1998) suggests that only extensive measure functions are admissible in pseudopartitives.

³⁵ Attested example, *Calgary Herald*, *Two degrees is all it takes—Warming may trigger floods* (December 17, 2009).

Unlike Lønning (1987), Schwarzschild (2006), and myself, Krifka's ontology does not have a layer of degrees that distinguishes between measure functions like *height* and unit functions like *meters*. Instead, he assumes that measure phrases involve functions that relate entities directly to numbers in a way that maps the ordering relation between entities "be smaller than" to the relation "be less than" between numbers. Thus, a function we might write *height-in-meters* relates entities directly to numbers that represent their height in meters, and similarly for functions like *volume-in-liters* and *temperature-in-degrees-Celsius*. I use these longer names only to distinguish the functions they denote from my measure and unit functions (see Section 2.5.4). Krifka (1998) refers to the functions he uses as *measure functions* and uses short names like *liters*. This analysis is advocated in Quine (1960: 244–5) and has been adopted in various other places, such as Chierchia (1998a: 74). As Schwarzschild (2002) points out, this analysis makes it difficult to analyze the semantics of comparatives like *six ounces heavier*. I adopt it only temporarily for the purpose of discussing the analysis of Krifka (1998). I will refer to the functions Krifka adopts as *Quinean measure functions*.

Here is how extensive measure functions are defined in Krifka (1989a, 1998). I have slightly condensed and adapted the proposal for presentation purposes. Here, \leq is mereological parthood, $+$ is arithmetic sum, and $>$ is the arithmetic greater-than relation.

(7) **Definition: Extensive measure function (Krifka)**

Let \wedge (a "concatenation") be an associative and commutative but non-idempotent operation. A measure function μ is extensive on \wedge iff for any a and b that are disjoint, $\mu(a) + \mu(b) = \mu(a \wedge b)$, and for any c and d , if $c \leq d$ and $\mu(d) > 0$ then $\mu(c) > 0$.

In order to avoid issues related to overlapping entities, Krifka's definition relies on the notion of concatenation. Krifka (1989a) is not very explicit about how his system selects a concatenation before checking if the measure function of a given pseudopartitive is extensive on it. Krifka (1998) suggests using mereological sum restricted to disjoint entities as a concatenation operation. Though this is a natural assumption, we will see that it leads to problems. But let us first look at a few cases in which the system works as expected.

A Quinean measure function like *volume-in-liters* is extensive in the sense of Krifka's definition in (7) because the volume of the sum of any two nonoverlapping entities is the sum of their volumes, and no entity with nonzero volume has a part with zero volume. (Such a part would have to be an empty part or bottom element. Krifka assumes a system equivalent to classical extensional mereology, in which there is no bottom element. See Section 2.3.1 for discussion.) A Quinean measure function like *temperature-in-degrees-Celsius* is not extensive because the temperature of the sum of two nonoverlapping entities is not the sum of their temperatures. Depending on the ontological setup, it might be undefined or their average.

Krifka places a constraint into the semantic translations of measure phrases like *thirty liters* that makes them compatible only with extensive Quinean measure functions in the sense of his definition. Schwarzschild (2006) notes that placing this constraint into the measure phrase causes problems because not all uses of measure phrases occur in constructions that reject intensive measure functions. Of course, Krifka's constraint could equally well be placed into the lexical entry of *of* in order to tie it to pseudopartitives specifically.

Turning now to Schwarzschild (2006), he uses the term *monotonic measure function* for a similar concept. While he does not provide a formal definition of monotonicity, the definition in (8) is implicit in the discussion in his paper. Roger Schwarzschild (p.c.) informs me that this definition is indeed what he had in mind.

(8) **Definition: monotonic measure function (Schwarzschild)**

A measure function μ is *monotonic* iff for any a, b , if a is a proper part of b , then $\mu(a) < \mu(b)$.

This definition is similar to the one in (7). This becomes clear when we apply it to the same example: volume is monotonic in the sense of this definition because any proper part of an entity has a smaller volume than that entity. But temperature is not monotonic because proper parts of an entity are generally not colder than that entity.

Both Krifka and Schwarzschild intend to draw a general connection between pseudopartitives and measure adverbials. Krifka (1998: sect. 3.4), notes that *for*-adverbials are like nominal measure phrases "insofar as they introduce a quantitative criterion of application." Schwarzschild (2006: sect. 3.2) points out that a formalization of the telic/atelic opposition in the line of Dowty (1979) can be couched in terms of monotonicity. For example, *in*-adverbials can only combine with telic predicates because, as he puts it, runtime "is nonmonotonic on the relevant part-whole relation in the domain given by" that predicate.

However, the actual definitions used by Krifka and Schwarzschild do not fit into the general picture pursued in this book. For one thing, they are not compatible with the behavior of distributive constructions. Through the Distributivity Constraint, adverbial-*each* constructions impose identical constraints on their Maps as pseudopartitives do (see Chapter 4). These thematic roles are generally not extensive or monotonic. Two nonoverlapping events, such as John's running from his house halfway to the store and his subsequent running to the store, can have the same agent, and their sum (John's running from his house to the store) again has the same agent.

In fairness, neither Krifka nor Schwarzschild attempted to relate pseudopartitives to adverbial-*each* constructions. But we find the same problem even if we only look at pseudopartitives. Height is acceptable in pseudopartitives like (9), so if Krifka's and Schwarzschild's accounts are correct, it should be extensive and monotonic.

(9) Five feet of snow covered Berlin.

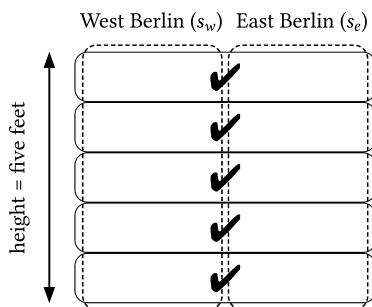


FIG. 7.1 The five feet of snow that fell on Berlin.

However, according to the definition underlying Schwarzschild (2006), height is not monotonic with respect to mereological parthood. Here is why: Imagine that it snows on Berlin for five days in a row and that the snow does not melt, so that there are then five layers of snow on top of each other. Assume that the height of the total snow cover is five feet. Let s be the snow that fell on Berlin. There are of course different ways of dividing s up. We can look at its horizontal layers, but we can also separate it vertically, according to the different regions on which it has fallen. For example, among the proper parts of s are the snow that fell on West Berlin, call it s_w , and the snow that fell on East Berlin, call it s_e . Then s is also the sum of s_w and s_e . This is illustrated in Figure 7.1. In line with the conventions in the figures in Chapter 6, the checkmarks serve as a reminder that each of the snow layers itself qualifies as snow. The problem is that s_w and s_e both have the same height. Each of them is five feet high. So height is not monotonic in the sense of Schwarzschild's definition.

Schwarzschild (2006) is aware of this problem. From similar examples, he concludes that pseudopartitives do not test for monotonicity with respect to the mereological part-whole relation, but with respect to a different part-whole relation which he sees as contextually supplied. In our example, his assumption would be that context provides a relation according to which the snow that fell on West Berlin, s_w , may well not be a part of the snow that fell on the entire city, s . Schwarzschild may well accept that s_w is a mereological part of s since snow is a mass noun, but this fact does not enter the picture.

I see two problems with this suggestion. First, Schwarzschild (2006) does not impose any formal constraints on the contextually supplied part-whole relation he assumes. We have already faced a similar situation in connection with the discussion of Moltmann's contextually determined part-whole relation in Section 5.3.3. At that time, I mentioned the objection by Zucchi & White (2001) to her proposal: "Since Moltmann does not tell us much about what *relevant* parts are, it is unclear to

what extent her formulation actually solves the minimal parts problem.” The same objection holds for Schwarzschild’s relation. There is no way to know whether two entities stand in Schwarzschild’s contextual part relation, so it is unclear how to test the predictions of his account. Second, many measure functions like *temperature* are already correctly ruled out even without replacing the mereological part-whole relation by a contextually supplied relation, so the two relations must coincide to a large extent.

A similar problem occurs with Krifka’s definition if we use mereological sum restricted to disjoint entities as our concatenation, as suggested in Krifka (1998). Height is not extensive on that concatenation. Since s_w and s_e are disjoint, their sum s is also their concatenation. The arithmetic sum of the height-in-feet of s_w and that of s_e is twice the height-in-feet of s rather than equal to it. The only kind of concatenation with respect to which height-in-feet would be extensive is one that requires a and b in (7) to range only over horizontal layers of snow that cover all of Berlin, in which case s_w and s_e will be excluded. Here we see the similarity to the strata that play a central role in this book. While there is nothing in Krifka’s account that would force this specific concatenation to apply, my own account can be seen as building it into the core semantics of the pseudopartitive construction.

7.4 Strata theory and measurement

In contrast to Schwarzschild (2006) but in keeping with Krifka (1998), my account is based on the mereological parthood relation. This relation is assumed to be independent of context (see Section 2.3.1). It is not necessary to appeal to context in order to account for the Berlin example, because stratified reference can already accommodate it. The issue and its solution are exactly the same as the solution discussed in Chapter 6 for the sentence *John pushed carts all the way to the store*. Instead of checking whether the Share predicate holds of every part of the entity in question, we only check whether there is a way to divide the entity into strata such that the Share holds of each stratum.

To draw on the intuition behind Krifka’s and Schwarzschild’s accounts without making an appeal to context necessary, I will relativize monotonicity to a certain dimension, such as height in the snow example. Furthermore, to account for the effect that the choice of substance noun matters, I will relativize it to the property denoted by that noun. I will do this by using stratified reference, which applies to a given property and entity, and which can be parametrized for a given dimension. In Chapter 6, I used the difference between temporal and spatial *for*-adverbials as motivation for the dimension parameter.

As we have seen, Schwarzschild suggests that the telic/atelic opposition can be formalized in terms of monotonicity. For example, *in*-adverbials can only combine

with telic predicates because, as he puts it, runtime “is nonmonotonic on the relevant part–whole relation in the domain given by” that predicate. For him, runtime is a dimension that is monotonic on the part–whole relation that relates events to their subevents. Here Schwarzschild uses monotonicity to subsume what I have captured through stratified reference. I propose to go in the opposite direction and use stratified reference to subsume what Schwarzschild captures through monotonicity.

Using stratified reference for both measure adverbials and pseudopartitives leads us to expect the link between the two domains that Krifka and Schwarzschild discuss. As we saw in Chapter 6, English has not only temporal but also spatial *for*-adverbials. The question arises whether other measure functions might also be acceptable in that construction. It appears that at least those measure functions that are rejected by pseudopartitives are also rejected by *for*-adverbials:

- | | | | |
|------|----|---|-----------------------|
| (10) | a. | John waited for five hours. | <i>temporal trace</i> |
| | b. | The crack widens for five meters. | <i>spatial extent</i> |
| | c. | *John drove for one ton. | * <i>weight</i> |
| | d. | *John drove for thirty miles an hour. | * <i>speed</i> |
| | e. | *The soup boiled for 100 degrees Celsius. | * <i>temperature</i> |

To some extent, these gaps have independent explanations. Unlike typical pseudopartitives, *for*-adverbials measure events rather than substances. Common sense suggests that driving events do not have weights, and we can therefore assume that examples like (10c) are category mistakes. More formally, I assume that the measure functions like *weight* are partial functions (see Section 2.5.3) that do not have any driving events in their domains. What is surprising, however, is that (10d) and (10e) are unacceptable with a *for*-adverbial but become acceptable when the word *for* is replaced by *at*, or in the case of (10d), left out:

- | | | | |
|------|----|---|--------------------|
| (11) | a. | John drove (at) thirty miles an hour. | <i>speed</i> |
| | b. | The soup boiled at 100 degrees Celsius. | <i>temperature</i> |

This contrast shows that we can talk in principle about the speed and temperature of an event, suggesting that examples like (10d) and (10e) cannot be ruled out on the grounds that events are not the kinds of things that have speeds and temperatures. The example from Davidson (1969), mentioned in Section 2.4.4, of a sphere rotating slowly and heating up quickly at the same time also suggests that speed is among the possible properties of events. We can avoid the undesirable conclusion that the sphere is both quick and slow by assuming that the rotating and the heating up are two separate events, and that each one has a different speed.

Just as we did in the case of pseudopartitives, we can identify properties which make it possible to use measure functions that are otherwise incompatible with *for*-adverbials. The following example shows this for the case of temperature:

- (12) The sample continued to cool for several degrees to point N and then suddenly increased to a temperature between the transition points of Form I and Form II with no indication of the presence of Form III.³⁶

These facts about *for*-adverbials provide further justification for the decision to link the domains of aspect and measurement via stratified reference.

Indeed, the account presented and motivated in the previous chapters explains these facts without any additional assumptions. Chapter 4, especially Section 4.3.2, has motivated the assumption that pseudopartitives are distributive constructions and that as such they obey the following constraint:

(13) **Distributivity Constraint**

A distributive construction whose Share is *S*, whose Map is *M*, and which is used to describe an entity *x* is acceptable only if *S* stratifies *x* with respect to *M* and a granularity level *g* specified by the construction (formally: $SR_{M,g}(S)(x)$).

Let us apply this constraint to the pseudopartitives in (4a) and (4b). The background assumptions described in Section 3.2 assign them the following LFs and logical representations (I have marked Key, Share, and Map for convenience):

- (14) a. $[[_{Key} \text{ thirty liters}] [[_{Map} \text{ volume}] [\text{of } [_{Share} \text{ water}]]]]$
 b. $[[_{Key} \text{ thirty degrees}] [[_{Map} \text{ temperature}] [\text{of } [_{Share} \text{ water}]]]]$
- (15) a. $\lambda x[\text{water}(x) \wedge \text{liters}(\text{volume}(x)) = 30]$
 b. $\lambda x[\text{water}(x) \wedge \text{degrees.celsius}(\text{temperature}(x)) = 30]$

I now show how the Distributivity Constraint in (13) accepts the representation in (15a) but rules out the representation in (15b). The abbreviation SR in (13) refers to stratified reference, whose definition is repeated here from Section 4.6:

(16) **Definition: Stratified reference**

Let *d* (a “dimension”) be any function from entities of type α to entities of type β , and let *g* (a “granularity level”) be any predicate of entities of type β . Let *P* range over predicates of type $\langle \alpha, t \rangle$ where α is either *e* or *v*, and let *x* range over entities of type α . Then:

$$SR_{d,g}(P)(x) \stackrel{\text{def}}{=} x \in * \lambda y \left(\begin{array}{c} P(y) \wedge \\ g(d(y)) \end{array} \right)$$

(A predicate *P* (the “Share”) stratifies *x* with respect to a function *d* (the “dimension” or “Map”) and a predicate *g* (the “granularity level”) iff *x* can be exhaustively divided into parts (“strata”) which are each in *P* and which are each mapped by *d* to something in *g*.)

³⁶ Attested example, Daubert & Clarke (1944).

As described in Chapter 4, I assume that pseudopartitives set g to $\gamma(\mu, x)$, a shorthand for $\lambda d.d < \mu(x)$, where μ is the Map and x is the entity (i.e. the substance or event) described by the pseudopartitive. I have encoded these assumptions in my lexical entry for *of*, repeated here:

$$(17) \llbracket \text{of} \rrbracket = \lambda S_{\langle \alpha, t \rangle} \lambda M_{\langle \alpha, \beta \rangle} \lambda K_{\langle \beta, t \rangle} \lambda x_{\alpha} : \text{SR}_{M, \gamma(M, x)}(S)(x) \cdot S(x) \wedge K(M(x))$$

When this is applied to the pseudopartitives in (4a) and (4b), stratified reference expands into the following conditions (details of the compositional derivation are given in Chapter 4). Here, x is the water amount that these pseudopartitives describe:

- (18) a. $x \in * \lambda y \left(\begin{array}{l} \text{water}(y) \wedge \\ \text{volume}(y) < \text{volume}(x) \end{array} \right)$
 (The entity x can be divided into parts, each of which is water and has a smaller volume than x does.)
- b. $x \in * \lambda y \left(\begin{array}{l} \text{water}(y) \wedge \\ \text{temperature}(y) < \text{temperature}(x) \end{array} \right)$
 (The entity x can be divided into parts, each of which is water and has a lower temperature than x does.)

As discussed in Section 2.6.5, given that *water* is a mass noun, it has approximate divisive reference. This means that every amount of water x apart from very small amounts can be divided into parts, each of which is itself water. Volume is extensive on water, that is, every part of any amount of water x has a smaller volume than x . Condition (18a) is therefore fulfilled, and (4a) is predicted to be acceptable. By contrast, temperature is not extensive on water. Consequently, condition (18b) is false, and (4b) is predicted to be unacceptable.

The Distributivity Constraint leads to the prediction that the same presupposition that is found in *for*-adverbials is also found in pseudopartitives. The intuition here is that a *for*-adverbial construction like *run for three hours* has essentially the same semantics as the corresponding pseudopartitive construction *three hours of running*, and that it gives rise to the same presupposition.

Given the Distributivity Constraint, examples like (10d) and (10e) can be ruled out because the presuppositions that stratified reference generates for them can plausibly be assumed to fail:

- (19) *drive for 30 miles per hour
Failed presupposition: $\text{SR}_{\text{speed}, \lambda d.d < \text{speed}(e)}(\text{drive})(e) \Leftrightarrow$
 $e \in * \lambda e' \left(\begin{array}{l} \text{drive}(e') \wedge \\ \text{speed}(e') < \text{speed}(e) \end{array} \right)$
 (The event e can be divided into parts, each of which is a driving event that is slower than e .)

- (20) *boil for 100 degrees Celsius

Failed presupposition: $SR_{temperature, \lambda d. d < temperature(e)}(boil)(e) \Leftrightarrow$

$$e \in * \lambda e' \left(\begin{array}{l} boil(e') \wedge \\ temperature(e') < temperature(e) \end{array} \right)$$

(The event e can be divided into parts, each of which is a boiling event that is colder than e .)

These presuppositions fail because the subevents of a driving event typically have the same speed as that event, and similarly for the subevents of boiling events and their temperatures. Even if one can imagine that some parts of a given driving event might be slower than the whole event, these parts cannot sum up to the whole event.

Since stratified reference is presupposed to hold of verb phrases, it is not surprising that specific verb phrases can rescue constructions that would otherwise be unacceptable, as in the case of temperature-based *for*-adverbials. For the following example, which is modeled on (12), I assume that the measure function relevant in (6), which I write *temperature-drop*, maps any cooling event to the number of degrees that it causes the temperature to drop.

- (21) cool for five degrees

Satisfied presupposition: $SR_{temperature-drop, \lambda d. d < temperature-drop(e)}(cool)(e) \Leftrightarrow$

$$e \in * \lambda e' \left(\begin{array}{l} cool(e') \wedge \\ temperature-drop(e') < temperature-drop(e) \end{array} \right)$$

(The event e can be divided into parts, each of which is a cooling event that involves a smaller temperature drop than that of e .)

On the plausible assumption that the cooling process in question causes the temperature of the affected entity to drop continuously, this presupposition is satisfied.

In substance-denoting pseudopartitives, I assume that the dimension parameter is set to the appropriate measure function. For example, in *thirty liters of water*, this measure function is volume, and the resulting presupposition is plausibly satisfied:

- (22) thirty liters of water

Satisfied presupposition: $SR_{volume, \lambda d. d < volume(x)}(water)(x) \Leftrightarrow$

$$x \in * \lambda y \left(\begin{array}{l} water(y) \wedge \\ volume(y) < volume(x) \end{array} \right)$$

(The entity x can be divided into parts, each of which is water whose volume is smaller than that of x .)

Mass nouns like *water* (as we have just seen) and plural count nouns like *books* are acceptable on the plausible assumption that they have approximate divisive reference (Link 1983, Krifka 1998): Whenever they apply to an entity, they also apply to all of its parts (leaving aside very small parts). By contrast, singular count nouns are ruled out on the assumption that they are quantized (i.e. they do not apply to any proper parts of any entity to which they apply), as proposed by Krifka (1998). This is shown in (23):

(23) *five pounds of book

Failed presupposition: $SR_{weight, \lambda d. d < weight(x)}(book)(x) \Leftrightarrow$

$$x \in *_{\lambda y} \left(\begin{array}{l} book(y) \wedge \\ weight(y) < weight(x) \end{array} \right)$$

(The entity x can be divided into parts, each of which is a book whose weight is less than that of x .)

The presupposition in (23) fails because a book does not consist of proper parts which are themselves books. The assumption that singular count nouns are quantized rules out the possibility that a book has any parts aside from itself, so it cannot have any parts whose weight is less than its own.

To be clear, it is not the stratified-reference presupposition that requires the entity x in (23) to be a book. That requirement comes from the at-issue meaning of the pseudopartitive. As described in Chapter 3, I assume following Schwarzschild (2006) and others that pseudopartitives are interpreted intersectively. That is, *five pounds of X* is true of any entity x that satisfies the predicate X and whose weight is five pounds. In (23), X is the singular count noun *book*. Therefore the predicate denoted by (23) will fail to hold of any entity x which is not a book. This is the reason why expressions such as *five pounds of book* are ruled out if *book* is interpreted as a count noun. The predicate denoted by this expression will fail to hold of any book because no book will satisfy the stratified-reference presupposition in (23). As for a sum of books whose joint weight is five pounds, even though it satisfies that presupposition, it will fail to satisfy the predicate denoted by (23) because it is not itself a book.

Turning to the problematic example of the snow that fell on Berlin, it can be given an account as follows:

(24) five feet of snow

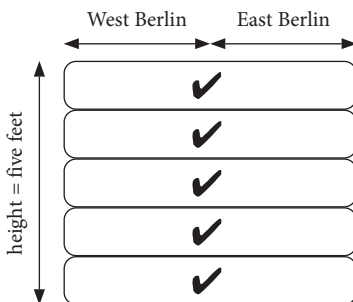
Satisfied presupposition: $SR_{height, \lambda d. d < height(x)}(snow)(x) \Leftrightarrow$

$$x \in *_{\lambda y} \left(\begin{array}{l} snow(y) \wedge \\ height(y) < height(x) \end{array} \right)$$

(The entity x can be divided into parts, each of which is an amount of snow whose height is less than that of x .)

This presupposition is satisfied in the Berlin scenario because horizontal layers of snow can play the role of the parts. The presupposition does not require height to be monotonic, and the example is acceptable in spite of the fact that the snow on West Berlin and the snow on East Berlin each have the same height as the snow that fell on Berlin as a whole.

The idea behind this account can again be understood via the visual metaphor introduced in Chapter 1. A plural or mass entity to which a pseudopartitive applies is divided into strata which are constrained as measured in the dimension determined by the pseudopartitive, but may extend arbitrarily in other dimensions. These strata are then required to be in the denotation of the noun. Singular count nouns always fail


 FIG. 7.2 Accepting *five feet of snow*.

this test because the individuals in their denotation are atomic, and cannot be further subdivided into strata. Figure 7.2 illustrates what (24) expresses formally: The measure function *height* is acceptable in the pseudopartitive *five feet of snow* because the snow cover in question can be divided into parts (horizontal layers) of snow whose height is less than five feet.

Like previous accounts, we still rule out examples involving temperature and similar intensive measure functions:

- (25) *thirty degrees Celsius of water

Failed presupposition: $SR_{temperature, \lambda d. d < temperature(x)}(water)(x) \Leftrightarrow$

$$x \in * \lambda y \left(\begin{array}{l} water(y) \wedge \\ temperature(y) < temperature(x) \end{array} \right)$$

(The entity x can be divided into parts, each of which is water which is colder than x .)

This presupposition fails as desired since the parts of a given amount of water will generally have the same temperature as the entire amount. Even if there are occasional fluctuations, it will not be possible to divide an amount of water into parts that will all have a lower temperature than the whole.

The constraint takes the substance noun into account, so we correctly predict the contrast between unacceptable cases like (4) (**five degrees (Celsius) of water*) and acceptable cases like (6) (*two degrees of global warming*). I assume that the relevant measure function, which I write *temperature-increase*, maps any warming event to the number of degrees of warming that it causes.

- (26) two degrees Celsius of global warming

Satisfied presupposition:

$SR_{temperature-increase, \lambda d. d < temperature-increase(x)}(global.warming)(x) \Leftrightarrow$

$$x \in * \lambda y \left(\begin{array}{l} global.warming(y) \wedge \\ temperature-increase(y) < temperature-increase(x) \end{array} \right)$$

(The entity x can be divided into parts, each of which is an amount of global warming that involves a smaller temperature increase than x does.)

The granularity parameter of stratified reference also makes it natural to account for an observation by Bale (2009): Pseudopartitives that are used to talk about very small quantities accept mass nouns but reject count nouns. A pseudopartitive cannot be predicated of an entity that is too small to be in the denotation of the substance noun. This is shown in (27).

- (27) a. Give me 500 grams of { apple / apples }.
b. Give me 100 grams of { apple / ?apples }.
c. Give me one gram of { apple / ??apples }.

As discussed earlier, pseudopartitives are incompatible with singular count nouns because these nouns are quantized and this leads to a failure of the stratified-reference presupposition. This means that whenever the word occurs in the singular, it is the mass sense of the word *apple* (which denotes portions of apple, i.e. applesauce or apple slices) which is at play here. As for *apples*, it denotes the plural of the count noun *apple* rather than of the mass noun *apple*, because most mass nouns can only be pluralized after being coerced to count nouns (see Section 2.6). The semantics of plural formation can be represented using the star operator (see Section 2.6.2). This means that portions of apple do not enter the extension of *apples* because they are not in the denotation of the singular count noun, and because the star operator closes predicates under sum but not under parthood. The status of examples (27b) and (27c) can be accounted for on the assumption that any two apples that the hearer could give the speaker will have a combined weight of over 100 grams but not necessarily over 500 grams.

The difference between mass and count uses of *apple* is also shown in the following attested example:

- (28) Lee said Americans eat the equivalent of one-fifth of a fresh apple each day, or about 19.7 pounds a year. But they should eat five times that much—at least one apple a day, he said. The Cornell researchers found that just 100 grams of apple have the same antioxidant activity as 1,500 milligrams of Vitamin C. (The average apple weighs 150 grams, or about 5 ounces.)³⁷

This example becomes worse if the substance noun in *100 grams of apple* is replaced by *apples*. Intuitively, this substitution has the effect that the example is now about whole apples instead of portions of apple. Stratified reference predicts this fact on the assumption that no two apples have a combined weight of as little as 100 grams, and that the weight of applesauce (i.e. of stuff that qualifies as *apple* in the mass sense) can be very small.

³⁷ Attested example, accessed May 25, 2016. <http://www.emphysema.net/Newsletters/Nov2004NL/November2004.htm>.

7.5 Summary

The starting point for this chapter was the old observation that intensive measure functions like *temperature* may not occur in pseudopartitives. We have seen that the same constraint is also operative in *for*-adverbials. Both facts are predicted by the present framework.

The constraint against intensive measure functions is also observed by Krifka (1998) and Schwarzschild (2006). While they also discuss some parallels between pseudopartitives and aspect, the present account is the first to explore the connections between different distributive constructions systematically. As we have already seen in Chapter 3, an event pseudopartitive like *three hours of running* is given the same analysis as a *for*-adverbial like *run for three hours*. This has provided us with the basis for explaining that the two constructions also license the same measure functions.

We can characterize the class of admissible measure functions as follows. A pseudopartitive has to satisfy stratified reference, where the dimension parameter is specified by the measure function and the granularity parameter is specified based on the measure phrase of the pseudopartitive. The constraint on measure functions is also instantiated in true partitives, comparative determiners, *for*-adverbials, and other constructions. An event pseudopartitive like *three hours of running* is given the same analysis as a *for*-adverbial like *run for three hours*. This explains why the two constructions also license the same measure functions. The constraint that rules out intensive measure functions like *temperature* proposed by Krifka (1998) and Schwarzschild (2006) is subsumed by the same constraint that also prevents telic predicates from combining with *for*-adverbials. The constraint against intensive measure functions does not have to be stipulated, because it is a consequence of the Distributivity Constraint. This helps us make sense of Krifka's and Schwarzschild's observations within the larger picture of strata theory.

In Chapters 5 and 6 and in this chapter, I exploited one of the defining features of strata theory, namely that it pushes us towards thinking of unboundedness as relativized to a certain dimension, thematic role, or measure function. For example, Chapter 6 argued that we should understand temporal atelicity as something more akin to the subinterval property than to divisive reference, because only the subinterval property is relativized to time. Metaphorically speaking, we should consider only the strata of a given event and not all its subregions. In this chapter, I have transferred this insight to pseudopartitives. While the entities involved are substances rather than events and while the dimensions are measure functions rather than thematic roles, the distinction between strata-based and subregion-based approaches is still operative. In this chapter, the example *five feet of snow* has played the same role as the example

push carts all the way to the store in Chapter 6. Both examples force us to consider two dimensions at once: height and width in the first case, time and space in the second. The insight from Schwarzschild (2006) that the constraint that comes with *five feet of snow* must be checked on horizontal layers of snow rather than on every subregion of snow finds a natural explanation in strata theory.

Covert distributivity

8.1 Introduction

This chapter presents a theory of covert distributivity that focuses on the distinction between lexical and phrasal distributivity and advocates a middle road in the debate on whether distributivity is atomic or nonatomic. I suggest a reformulation of the classical definition of distributivity operators and propose to expand their use into the temporal domain. The purpose of this chapter and the next is to bring together several strands of research on semantic and pragmatic phenomena in order to provide a comprehensive analysis of distributivity in natural language. This chapter has been previously published as Champollion (2016a), and it is reprinted here with slight modifications.

Covert distributivity as I will talk about it in this chapter is diagnosed by the ability of indefinites or numerals in object position to covary in a way that is generally attributed to a covert verb-phrase-level modifier called the D operator (Link 1987b, 1991a; Roberts 1987). The meaning of this kind of operator is either similar to the adverbial modifier *each* or it corresponds to something like *each salient part of*, where salience is a context-dependent notion (Schwarzschild 1996). For overviews of the major empirical phenomena related to distributivity, see also Section 4.2 here, Champollion (to appear: sect. 2), Lasersohn (2011), Nouwen (2016), Szabolcsi (2010: chs 7 and 8), and Winter & Scha (2015).

The ambiguity between distributive and scopeless readings in English can be modeled by assuming that the D operator is optionally present as a silent verb phrase modifier whose syntax and meaning correspond to that of adverbial *each*. For example, (1a) represents a scopeless reading and (1b) a distributive reading. I use the term *scopeless* to refer both to collective and cumulative readings. The distinction between these two readings does not matter for this chapter. See Landman (2000) and Section 2.8 for discussion.

- (1) a. The boys saw two monkeys.
 ≈ The boys between them saw two monkeys. *scopeless*
- b. The boys [D [saw two monkeys]].
 ≈ The boys each saw two monkeys. *distributive*

This chapter makes both technical and empirical contributions to semantic theory. The main technical contribution of this chapter is a reformulation of distributivity operators that makes them compatible with Neo-Davidsonian algebraic event semantics. This reformulation makes it possible to draw on the resources of event semantics in order to formally model the relations between distributivity over individuals and over events and times in a parallel way. Moreover, it makes it possible to draw on the resources of algebraic semantics and mereology in order to formally model the relations between distributivity over atoms and over nonatomic entities like time intervals in a parallel way.

The main empirical contribution of this chapter is a unified theory of covert atomic and nonatomic distributivity, over individuals and over temporal intervals, at the lexical and at the phrasal level. As discussed in Chapter 4, I understand distributivity as involving the application of a predicate to the members or subsets of a set, or to the parts of an entity (individual, event, or interval). This application is diagnosed by the presence of what I have called *distributive entailments*.

There has been a longstanding debate about whether and to what extent distributivity over nonatomic entities, or “genuine plural quantification” (Link 1987b), ever occurs (Link 1987b; Gillon 1987, 1990; Lasersohn 1989; Schwarzschild 1996; Winter 2001; Kratzer 2007). Here is a preview of the argument developed in Schwarzschild (1996). Schwarzschild argues that the distributivity operator should be modified to allow for nonatomic interpretations in a limited set of circumstances, essentially whenever there is a particularly salient way to divide a plural individual into parts other than its atoms. Here is an example. Shoes typically come in pairs, so a sentence like (2) can be interpreted as saying that each pair of shoes costs fifty dollars, as opposed to each shoe or all the shoes together.

- (2) (Context: 3 pairs of shoes are on display, each pair with a \$50 tag)

The shoes cost fifty dollars. (Lasersohn 1998b)

Since the numeral *fifty dollars* covaries with pairs of shoes rather than with shoes, this example is generally taken to involve nonatomic distributivity. The presence of this kind of interpretation depends on contextual factors. For example, it is not part of the meaning of sentence (2) itself that shoes typically come in pairs. Verb phrases can only be interpreted as distributing over nonatomic entities if there is supporting context or world knowledge that makes these nonatomic entities pragmatically salient. In the absence of this support, verb phrases must distribute over atoms or not at all, as the following example shows.

- (3) John, Mary, Bill, and Sue were paid fifty dollars. (based on Lasersohn 1989)

This example can be interpreted as saying that the four people in question were each paid fifty dollars, or that they were paid fifty dollars together. Out of the blue, nonatomic interpretations are not available. For example, in a scenario where each of

the four people in question was paid twenty-five dollars, the sentence is false, even though there are ways to group the four people into pairs such that each pair was paid a total of fifty dollars.

Schwarzschild (1996) suggests that the difference between (2) and (3) is due to the lack of a contextually salient partition or cover in the latter case. He models this by making the distributivity operator anaphoric on such a cover, and renaming it the Part operator. In this chapter, I adopt this strategy and extend it to the temporal domain, to create a parallel between the discussion of nominal and temporal nonatomic distributivity. The parallel can be illustrated by the following pair of examples:

- (4) a. John found a flea { for ten minutes / for a month }.
- b. The patient took two pills for a month and then went back to one pill.

The ten-minute variant of (4a) is from Zucchi & White (2001), while (4b) is based on observations in Moltmann (1991). Out of the blue, examples like these two are odd, and seem to suggest that the same flea is found repeatedly and the same pills are taken repeatedly. Thus, Zucchi & White give their example two question marks, and describe it as not acceptable unless understood iteratively. Iterative interpretations can improve such examples, but they require special contexts. At least in the case of (4b), it is easy to find such a context. For example, (4b) is acceptable in a hospital context where the patient's daily intake is salient. It does not require any pill to be taken more than once. In the case of (4a), finding such a context is harder, particularly for the ten-minute variant.

I will argue that this diagnoses verb-phrase-level distributivity, and that it is caused by the presence of a temporal analogue of the Part operator discussed above. The contribution of this operator in example (4b) can be paraphrased as *every day*. The operator takes scope over the verb phrase *take two pills* but under the *for*-adverbial. The month-long interval introduced by the *for*-adverbial plays the same role vis-à-vis the distributivity operator as the collective individual denoted by *the shoes* does in example (2).

I will contrast this theory with an alternative view, on which the covariation in (4b) is due to the *for*-adverbial itself, and no distributivity operator is present in examples like it. On that alternative view, the *for*-adverbial is interpreted as a universal quantifier meaning something like *at each relevant point during a month*, as has been suggested at various times in the literature (Dowty 1979; Moltmann 1991; Deo & Piñango 2011). I will argue that *for*-adverbials cannot be interpreted as universal quantifiers, since out of the blue they do not induce covariation in indefinites they outscope—see (4a)—except when an overt distributive quantifier intervenes. This is shown in example (5), adapted from Zucchi & White (2001), where a multiple-fleas interpretation is readily available:

- (5) John found a flea on his dog every day for a month.

In examples where the “same-object” and “different-objects” interpretations are both plausible, a similar contrast can be observed. In (6a), one golf ball must have been hit and retrieved repeatedly, while in (6b) it is also possible that each time a different ball was hit (Sandro Zucchi p.c. to van Geenhoven 2005):

- (6) a. Jim hit a golf ball into the lake for an hour.
- b. Jim hit a golf ball into the lake every five minutes for an hour.

I will account for the contrast between (4a) and (5), and analogously for the contrast between (6a) and (6b), by claiming that the temporal distributivity operator cannot occur in the former, and that *every day* or *every five minutes* plays its role in the latter. This requires an explanation of why the operator is not able to occur in (4a) or (6a) out of the blue. I will suggest that its (in)ability to occur is due to a contextual factor: It is anaphoric on a salient set of stretches of spacetime, in the same way as the nonatomic distributivity operator in (2) is anaphoric on a salient set of shoes.

While the topic of this chapter is covert distributivity, Chapter 9 is about overt distributivity, as manifested in adverbial *each* and its adnominal and determiner counterparts both in English and other languages (Zimmermann 2002b). The meanings of these items varies in ways that sometimes require them to distribute over individuals (such as in the case of *each*) and in other cases allow them to distribute over salient parts of spacetime (such as in the case of German *jeweils*). The main claim of Chapter 9 is that the D operator relates to the Part operator in the same way as *each* relates to *jeweils*. Thus, overt and covert distributivity share many similarities. This gives rise to similar questions in the two cases. Can a distributivity operator only distribute down to singular entities or also to plural entities? Do these entities need to be of a certain size or “granularity,” and can this size vary from operator to operator? Must these entities have been overtly mentioned in the sentence and thereby contributed by semantic means, or can they also be supplied by the context via pragmatic means?

A unified semantic analysis of distributivity should make it apparent which aspects of the meanings of various distributivity operators are always the same, and along which dimensions these meanings can differ. I will explain the fact that the various overt and covert distributivity operators share some part of their meanings. To do so, I will adopt the strata-theoretic view of distributivity as the property of a predicate which, whenever it holds of a certain entity or event, also holds of its parts along a certain dimension and down to a certain granularity. As we have seen throughout this book, strata theory conceptualizes dimension and granularity as parameters which can be set to different values for different instances of distributivity. The dimension parameter takes a (partial) function as its value. It specifies the domain in which the predicate in question is distributed. For the purpose of this chapter, different settings of this parameter will allow us to capture the commonalities and differences between distributivity over discrete (count) domains and distributivity over domains involving continuous dimensions, such as time and mass domains. The granularity parameter

takes a predicate as its value. It is used to specify that the parts in question must be atomic, or that they can be nonatomic but must be constrained as measured along the specified dimension. This parameter accounts for the differences between distributive constructions over discrete and continuous domains. The two parameters interact with each other against the background of assumptions about the metaphysics of natural language. For example, I assume that time is nonatomic, or in any case that it does not make its atoms available to the semantics of natural language. As a result, when the dimension parameter is set to time, the granularity parameter cannot be set to *Atom*, because time does not provide any atoms to distribute over.

As will become clear, this understanding of distributivity provides several theoretical advantages. First, by understanding distributivity as parametrized for granularity, we gain a new perspective on the debate between proponents of atomic and cover-based formulations of distributivity operators. The atomic distributivity operator of Link (1987b), Roberts (1987), and Winter (2001) corresponds to one setting of the granularity parameter, and the nonatomic distributivity operator of Schwarzschild (1996) corresponds to another setting. Following Schwarzschild, I will assume that there is a distributivity operator whose granularity parameter is anaphoric on its context and can only be set to a nonatomic value when context supports a salient granularity level.

Second, by understanding distributivity as parametrized for dimension, we gain the technical ability to distinguish agent-based from theme-based distributivity and the like (Lasnik 1998b). Indeed, not only thematic roles like *agent* and *theme* can be considered dimensions, but also trace functions like *runtime* and *location*. We can therefore instantiate the dimension parameter of distributivity with time, or more specifically, with the temporal trace function τ that maps events to their runtimes (see Section 2.5.2). Given the assumption that time and space are nonatomic, we expect that this should only be possible when the granularity parameter of the distributivity operator is set to a nonatomic value, which in turn should require context to provide a salient granularity. I will argue that such contexts indeed exist, although they are rare. One example is the hospital context in which (4b) was to be understood. I show that the corresponding phenomenon has already been noticed in the literature on aspect. Through parametrized distributivity, the asymmetry between the atomic domain of individuals and the nonatomic domain of time allows us to explain the scopal behavior of *for*-adverbials. If the distributivity operator is easily available only when its granularity is atomic, then it is expected not to be easily available in the temporal domain of *for*-adverbials.

The Neo-Davidsonian event-semantic setting also gives us the ability to think of overt and covert distributivity operators as being (co-)indexable with different thematic roles. This allows us to capture through a simple change in indexation the kinds of configurations that have otherwise been taken to require type-shifting-based reformulations of the D operator (Lasnik 1998b):

- (7) a. The first-year students [D [took an exam]]. *Target: agent*
b. John [D [gave a pumpkin pie]] to two girls. *Target: goal*

Chapter 9 will show that this phenomenon has a direct counterpart in examples like (8) and (9), which involve adnominal *each* (Zimmermann 2002b; Blaheta 2003):

- (8) The boys told the girls two stories each. *Target: agent*
(two stories per boy)
(9) The boys told the girls two stories each. *Target: goal*
(two stories per girl)

The theoretical picture that is developed here provides a way to formalize such parallels across instances of distributivity in natural language. Individual items can be analyzed as being hardwired for certain parameter values, so that, for example, the difference between Link's and Schwarzschild's operators, as well as that between *each* and *jewels*, can be described in terms of whether the value of the granularity parameter is prespecified to *Atom* or can be filled in by context. In this way, overt and covert instances of distributivity fit together and into distributivity theory more generally.

The rest of this chapter is organized as follows. In Section 8.2, I review the literature on covert distributivity in the nominal domain, focusing on the atomic distributivity operator introduced in Link (1987b) and Roberts (1987). Nonatomic distributivity is discussed in Section 8.3, where I present the nonatomic distributivity operator introduced in Schwarzschild (1996), summarize the literature on the topic, and describe my own view. The question of how to adapt distributivity operators into event-semantic frameworks, such as the one adopted here, is discussed in Section 8.4. In that section, I argue for a specific way to do so, which I argue in Section 8.5 to be superior to a previous proposal by Lasersohn (1998b). I outfit the D and Part operators with dimension and granularity parameters. When the dimension parameter of the reformulated Part operator is set to time, the result induces covariation of indefinites over salient stretches of time. Section 8.6 builds on this result and provides an account of the limited ability of covariation by indefinites in the scope of *for*-adverbials. Section 8.7 summarizes and concludes.

8.2 Atomic distributivity

Let me now present the theory of atomic distributivity initiated by Link (1987b) and Roberts (1987), and further developed by subsequent authors. I start here with the atomic distributivity operator D as originally defined by Link and Roberts, and the motivation that led to it. In the next section, I focus on the nonatomic distributivity operator Part as originally defined by Schwarzschild (1996). The discussion here takes inspiration from Winter (2001: ch. 6), which in turn builds on Link (1987b), Roberts

(1987), and others. For other introductions to the same topic, see also Schwarzschild (1996: ch. 6), Link (1997: sect. 7.4), and Champollion (to appear: sect. 2).

Distributivity can be understood, among other things, as a property of predicates (see Section 4.2.3). Certain lexical predicates (i.e. predicates that consist of just one word) such as *smile*, and certain phrasal predicates (i.e. predicates that consist of several words, typically a verb and an object) such as *wear a black dress*, are distributive: whenever several people smile or wear a black dress, this entails that each of them smiles or wears a black dress. The distinction between lexical and phrasal predicates that have distributive interpretations will be important throughout this chapter. To highlight this distinction, from now on I will speak of lexical distributivity when there is a single word that is understood distributively, and I will speak of phrasal distributivity when a phrase that consists of more than one word is understood distributively. As the following examples show, we find distributive and collective interpretations in both classes of predicates:

(10) **Lexical distributivity/collectivity**

- | | |
|--------------------------------|---------------------|
| a. The children smiled. | <i>distributive</i> |
| b. The children were numerous. | <i>collective</i> |

(11) **Phrasal distributivity/collectivity**

- | | |
|---|---------------------|
| a. The girls are wearing a black dress. | <i>distributive</i> |
| b. The girls are sharing a pizza. | <i>collective</i> |

The classification of these interpretations as distributive and collective is based on the presence or absence of distributive entailments (see Section 4.2). Sentence (10a) entails that each child smiled, while sentence (10b) does not entail that each child was numerous. Similarly, sentence (11a) entails that each girl wears a different dress, but sentence (11b) does not entail that the girls ate different pizzas.

The distinction between lexical and phrasal distributivity is similar to the P/Q-distributivity distinction introduced in Winter (1997, 2001). Winter uses the term *P-distributivity* (where *P* stands for *predicate*) to refer to those cases of distributivity which can, in principle, be derived from some property of the verb involved. *Q-distributivity* (*Q* for *quantificational*) refers to cases where this approach is not possible because the distributive predicate contains an indefinite or numeral quantifier, as in (11a). In order for (11a) to entail that each girl wears a different dress, the entire verb phrase, including its object, must be distributed over the girls. This means that the entire verb phrase *wear a black dress* and not just the verb *wear* must be regarded as distributive. In this verb phrase, the quantifier is introduced by a separate word, so this is a case of phrasal distributivity. The difference between lexical and phrasal distributivity corresponds to the difference between what can and what cannot be ascribed to the lexical semantics of the verb. It is possible to ascribe the difference between (10a) and (10b) to the meaning of *smile* and *be numerous*. The difference

between the distributive interpretation in (11a) and the collective interpretation in (11b) is of a different kind, since it involves a scopal ambiguity. Accounts that are based only on lexical semantics cannot model Q-distributivity, and therefore cannot model phrasal distributivity, because they cannot create a scopal dependency between the definite subject and the indefinite object. For more details, see Winter (2001: sects 3.2 and 6.1).

The usual way to model Q-distributivity is to introduce a covert distributive operator in the logical representation so that the indefinite can take scope at two different places with respect to it. This is the purpose of the D operator postulated by Link (1987b) and Roberts (1987). It shifts a verb phrase to a distributive interpretation, more specifically, one that holds of any individual whose atomic parts each satisfy the unshifted verb phrase.³⁸ This D operator is usually defined as follows.

(12) **Link's D operator**

$$[[D]] \stackrel{\text{def}}{=} \lambda P \lambda x \forall y [y \leq x \wedge \text{Atom}(y) \rightarrow P(y)]$$

(Takes a predicate *P* over individuals and returns a predicate that applies to any individual whose atomic parts each satisfy *P*.)

Here, the variable *x* is resolved to a plural entity, typically provided by the subject, and *y* ranges over its atomic parts, that is, those parts of *x* which have only themselves as parts. These atomic parts intuitively stand for the singular individuals of which *x* consists. The operator introduces a universal quantifier, and it is the scopal interaction of this quantifier with the indefinite inside a Q-distributive predicate that accounts for the covariation effects. For example, if the verb phrase *wear a black dress* is represented as $\lambda x \exists z [\text{black}(z) \wedge \text{dress}(z) \wedge \text{wear}(x, z)]$, the meaning of (11a) can be represented in a way that places it in the scope of the universal quantifier introduced by the D operator:

- (13) $\forall y [y \leq_{\text{Atom}} \bigoplus \text{girl} \rightarrow \exists z [\text{black}(z) \wedge \text{dress}(z) \wedge \text{wear}(y, z)]]$
(Every atomic part of the sum of all girls wears a black dress.)

The D operator is able to apply to entire verb phrases and not just to lexical predicates. It is this property that allows it to account for phrasal distributivity (Dowty 1987; Roberts 1987; Lasersohn 1995). Moreover, at least Roberts (1987) allows the D operator to apply to any predicate, whether it is a verb phrase or not. For example, it may apply to a predicate that has been derived by λ -abstraction over a nonsubject predicate in order to derive an interpretation of (14) where each of two girls received a pumpkin pie:

- (14) John gave a pumpkin pie to two girls.
(two girls) $D[\lambda x. \text{John gave a pumpkin pie to } x]$ (Roberts 1987)

³⁸ The operator goes back to Link (1991a), which was written before Link (1987b). See Roberts (1987: 157) for discussion.

This approach involves an otherwise unmotivated structure or perhaps an application of quantifier raising, and is criticized for this reason by Lasersohn (1998b), whose own proposal is the topic of Section 8.5. The need for the D operator to be able to target noun phrases other than the subject is an important point, and I return to it in Section 8.4. As we will see, my own implementation deals with nonsubject predicates by parametrizing the D operator for different thematic roles. But first, I turn to a review of nonatomic distributivity.

8.3 Nonatomic distributivity

In Section 8.2, I presented the atomic view on distributivity. This view assumes that phrasal distributivity involves universal quantification over atomic parts of the plural individual, that is, over singular individuals. On this view, the distributive reading of a sentence like *The girls are wearing a black dress* is equivalent to *The girls are each wearing a black dress*. The indefinite *a black dress* covaries with respect to a covert universal quantifier that ranges over individual girls. This view is defended in Lasersohn (1995, 1998b), Link (1997), and Winter (2001), among others. By contrast, the nonatomic view holds that phrasal distributivity involves universal quantification over certain parts of the plural individual, and that these parts can be nonatomic. Variants of this view are defended in Gillon (1987, 1990), Verkuyl & van der Does (1991), van der Does & Verkuyl (1996), Schwarzschild (1996: ch. 5), Brisson (1998, 2003), and Malamud (2006a, 2006b). This section presents and motivates the nonatomic view.

Section 8.1 mentioned examples like (2) (*The shoes cost fifty dollars*), where context can make nonatomic interpretations available. This kind of example is discussed in Schwarzschild (1996). There, Link's distributivity operator is modified so that it is no longer restricted to distribute over atoms. The nonatomic view is based on sentences like the following, which is adapted from Gillon (1987):

- (15) Rodgers, Hammerstein, and Hart wrote musicals.

This sentence plays on a particular fact of American culture: Neither did the three composers it mentions ever write any musical together, nor did any of them ever write one all by himself. However, Rodgers and Hammerstein wrote the musical *Oklahoma!* together, and Rodgers and Hart wrote the musical *On Your Toes* together. On the basis of these facts, the sentence is judged true in the actual world, although it is neither true on the collective interpretation (since no musical was written by the three composers together) nor on an "atomically distributive" interpretation (since it is not true that each of them wrote any musical by himself).

An early argument for the nonatomic view was given as follows (Gillon 1987, 1990). In order to generate the reading on which (15) is true, the predicate *wrote musicals* must be interpreted as applying to nonatomic parts of the sum individual to which the subject refers. This view is generally implemented with the concept of a cover.

Recall from Section 5.4 that a set C is a cover of a plural entity x iff C is a set of parts of x whose sum is x .

Cover-based approaches modify the distributivity operator by relaxing the “atomic part” condition and by quantifying over nonatomic parts of a cover of the plural individual (Schwarzschild 1996). The first cover-based approaches assumed that the cover can be existentially quantified by the operator that introduces it. In an eventless setting, this assumption can be implemented by a distributivity operator such as (16). On this view, the denotation of sentence (15) can be represented as (17). This formula is verified in the actual world by the existence of the cover $\{\text{rodgers} \oplus \text{hammerstein}, \text{rodgers} \oplus \text{hart}\}$. The condition $y \leq x$ in (16) is redundant given that C is required to be a cover of x ; it is included only for clarity.

(16) **Nonatomic distributivity operator, existentially bound cover**

$$\llbracket D_{\exists} \rrbracket \stackrel{\text{def}}{=} \lambda P \lambda x \exists C [\text{Cov}(C, x) \wedge \forall y [C(y) \wedge y \leq x \rightarrow P(y)]]$$

$$(17) \quad \exists C [\text{Cov}(C, \text{rodgers} \oplus \text{hammerstein} \oplus \text{hart}) \wedge \forall y [C(y) \wedge y \leq x \rightarrow y \in \llbracket \text{wrote musicals} \rrbracket]]$$

Existentially bound covers are not a good way to model phrasal distributivity, because they overgenerate readings. These can be described as halfway between collective (or cumulative) and distributive readings, and they are sometimes called intermediate readings. I call them *nonatomic readings*. For example, suppose that John, Mary, Bill, and Sue are the teaching assistants, that each of them taught a recitation section, and that each of them was paid \$7,000 last year. Then sentences (18a) and (18b) are both true (Lasersohn 1989). This is as is expected on the atomic approach. Sentence (18a) is true on its distributive reading, and sentence (18b) is true on its collective or cumulative reading. But sentence (18c) is false, even though the cover $\{j \oplus m, b \oplus s\}$ would verify it if it was represented using the D_{\exists} operator in (16). That is, sentence (18c) does not have a nonatomic reading.³⁹

- | | |
|--|--------------------------------|
| (18) a. The TAs were paid exactly \$7,000 last year. | <i>atomically distributive</i> |
| b. The TAs were paid exactly \$28,000 last year. | <i>collective</i> |
| c. The TAs were paid exactly \$14,000 last year. | <i>*nonatomic</i> |

Giving up the existential cover-based operator D_{\exists} in (16) explains why sentence (18c) is false, because without this operator, there is no way to generate a nonatomic reading for this sentence. However, sentence (15) does have a nonatomic reading, so giving up D_{\exists} requires an alternative account of this reading and the inferences we can draw from it. Here it is important to note that not every inference requires an operator to account for it. For example, the inference from “The TAs have money” to

³⁹ While Lasersohn’s scenario involved only three individuals, I have added a fourth, in order to sidestep the question of whether nonatomic readings allow for overlap.

“Each TA has money” does not require a nonatomic distributivity operator because it follows from the assumptions that *have* is distributive and *money* has divisive reference (see Section 2.3.5).

Lasersohn (1989) proposes to account for the nonatomic reading of (15) through the use of lexical meaning postulates like (19):

- (19) $\forall x_1 x_2 y_1 y_2 [(\text{write}(x_1, y_1) \wedge \text{write}(x_2, y_2)) \rightarrow \text{write}(x_1 \oplus x_2, y_1 \oplus y_2)]$
 (Whenever x_1 writes y_1 and x_2 writes y_2 , x_1 and x_2 write y_1 and y_2 .)

This meaning postulate embodies the lexical cumulativity assumption (see Section 2.7.2). Cumulativity is assumed to be a property of all verbs, but not of all verbal constituents. Lexical meaning postulates like (19) differ from distributivity operators in that they are taken to apply only to verbs, and not to verb phrases or larger constituents. Following Kratzer (2007), and contra Sternefeld (1998) and Beck & Sauerland (2000), I will not generalize the cumulativity assumption from verbs to arbitrary constituents, as this would make it difficult to model the difference between lexical and phrasal distributivity.⁴⁰

Lasersohn’s meaning postulate ensures that when a cumulative verb, such as *write*, combines with a cumulative object, such as *musicals*, the result is a cumulative verb phrase. Thus when there are two entities—individual people or sums—that each qualify as *write musicals*, so does their sum. Given the inclusive view of the plural (see Section 2.6.2), *write musicals* literally applies to entities who wrote one or more musicals (Krifka 1989b; Sauerland, Anderssen & Yatsushiro 2005; Spector 2007). Such entities include the sum *rodgers* \oplus *hammerstein* and the sum *rodgers* \oplus *hart*, and via (19), the sum individual *rodgers* \oplus *hammerstein* \oplus *hart*, of which (19) entails that it wrote the sum individual *oklahoma* \oplus *on.your.toes*. This plural individual qualifies as *musicals*.⁴¹

Even with the meaning postulate in place, a noncumulative object that combines with a cumulative verb will not generally yield a cumulative verb phrase. Thus, the two sums *rodgers* \oplus *hammerstein* and *rodgers* \oplus *hart* each satisfy the truth conditions of *write a musical*, but in the absence of a cover-based distributivity operator, their sum does not. Indeed, the following sentence, unlike (15), is false in the actual world (Link 1997):

⁴⁰ In fact, it may even be necessary to prevent cumulativity from freely applying at the level of the verb. This free application amounts to the assumption that iterative interpretations come for free at that level (Kratzer 2007). By various measures, iterative interpretations lead to higher processing costs even at the level of the verb (Deo & Piñango 2011). If verb-level iterative interpretations are due to lexical cumulativity, these costs are unexpected; for relevant discussion, see Champollion (2013). In this chapter, I set aside the question of how to account for higher processing costs of iterative interpretations at the verb level. In Section 8.6, I briefly discuss a different kind of processing cost that occurs at the level of the verb phrase.

⁴¹ Sentence (15) conveys that more than one musical in total is written. This can be explained e.g. by modeling this information as a scalar implicature (Spector 2007; Zweig 2008, 2009). See Section 2.6.2.

(20) Rodgers, Hammerstein, and Hart wrote a musical.

If a cover-based operator like (16) was available in the grammar, that operator would predict (20) to be true in the actual world. Lasersohn and many others conclude from this and similar examples that the atomic approach to phrasal distributivity is superior to covers (e.g. Winter 2001).

However, some sentences do require cover-based operators (Gillon 1990; Schwarzschild 1996). They typically involve special contexts in which a specific cover is salient. This is where the shoe example in (2) comes in, repeated here:

(21) The shoes cost fifty dollars. (Lasersohn 1998b)

This sentence can be interpreted with respect to a cover whose cells each contain a matching pair of shoes. The relevant reading is nonatomic because it asserts that each pair of shoes costs fifty dollars and these pairs are nonatomic parts of the denotation of *The shoes*.⁴² The reading is not atomically distributive because it does not assert that each shoe costs fifty dollars, and it is not collective because it does not assert that all the shoes taken together cost that much.

Justin Bledin (p.c.) offers a similar example, but unlike (21), the salient cover in this case has overlapping cells. The example involves a magic square, a type of puzzle that involves filling in a grid with natural numbers so that the rows, columns, and diagonals all sum to the same thing. Suppose I give you a magic square to solve and I give you the following clue:

(22) The numbers sum to twenty-five.

The salient reading of this sentence is not atomically distributive because it does not assert that each number sums to twenty-five, and it is not collective because it is each row, column, and diagonal that sums to twenty-five, rather than the totality of the numbers.

By contrast, no cover is salient for example (23), and so it can only mean that each suitcase weighs fifty pounds or all of them together do so.

(23) The suitcases weigh fifty pounds.

In the nonatomic reading of (21) above, the quantifier introduced by the object takes scope under the distributivity operator. Unlike the relevant reading of sentence (15), this scopal dependency cannot be captured by a meaning postulate like (19). Schwarzschild (1996) models the context dependency of this kind of reading by assuming that the distributivity operator (which he renames *Part*, to set it apart from

⁴² I assume that pairs of shoes only occur in the denotation of *shoes* as proper sums, not as atoms, because *This is a shoe* cannot be uttered to describe a pair of shoes, and because somebody who owns one pair of shoes cannot answer *How many shoes do you own?* by *One*. An interesting question is whether *The shoes cost fifty dollars each* is an acceptable way to describe the relevant scenario.

Link's D operator) contains a free cover variable whose value is supplied anaphorically by context. I will represent this variable as a subscripted C .

(24) **Schwarzschild's nonatomic distributivity operator, free cover**

$$\llbracket \text{Part}_C \rrbracket \stackrel{\text{def}}{=} \lambda P \lambda x \forall y [C(y) \wedge y \leq x \rightarrow P(y)]$$

The difference between D and Part amounts to a division of labor between semantics and pragmatics. Semantics accounts for atomic phrasal distributivity, and pragmatics for nonatomic phrasal distributivity. Schwarzschild assumes that C is restricted through a pragmatic mechanism to be a cover over x , but he prefers not to write this condition into his operator.

The introduction of a pragmatic component into the analysis of what had previously been treated as a purely semantic phenomenon is discussed and justified at length in Schwarzschild (1996). Relevant evidence comes from sentences like (25):

- (25) The young animals and the old animals were separated.

(Schwarzschild 1996: 44)

This sentence typically entails that the young animals were separated from the old animals but that each of these two groups stayed together. This suggests that the verb phrase *be separated* is distributed down to the level of these two groups and not all the way down to individual animals. At this point, a proponent of atomic distributivity might argue that the reason that the verb phrase is able to apply at this intermediate level is that the two groups of animals are in fact atoms. The two conjuncts might then be analyzed as involving group-forming operators that map each of the two pluralities of animals to an atom. This route is taken by Landman (1989). Schwarzschild rejects this approach and argues for the essentially pragmatic nature of nonatomic distributivity by pointing out that the inference down to groups is cancelable. The following sentence leaves it open exactly how the animals were separated, a fact that is unexpected on the group-based analysis:

- (26) The young animals and the old animals were separated, but not necessarily by age.

Schwarzschild assumes that the *but*-clause in (26) prevents the value of C from being set to the cover that is made pragmatically salient by the conjunction. Beyond this kind of case, Schwarzschild does not say much about the precise pragmatic mechanism that resolves C . For a proposal in which the Part operator is anaphoric on a decision problem in the sense of van Rooij (2003), see Malamud (2006a, 2006b, 2012). From the present perspective, what matters most in Schwarzschild's and Malamud's approaches is that the Part operator imposes a stronger restriction on the identity of the cover than would be achieved by just existentially quantifying over it. This restriction rules out nonatomic readings in sentences like (18c), (20), and (23), but not in sentences like (21). While I see no obstacles to using decision problems,

I will continue to use a Schwarzschild-style approach to keep the representation simple.

Sentence (21) is structurally equivalent to sentences (18c), (20), and (23), yet only (21) has a nonatomic reading. As Heim (1994) and Schwarzschild (1996) argue, this fact provides strong evidence that models of (phrasal) distributivity need to contain a pragmatic factor. The operator in (24) is more restricted than the existential cover-based operator D_3 in (16) because (24) presupposes the existence of a context through which the variable C can be resolved. The contrast between (21), which has a nonatomic reading, and (18c), (20), and (23), which do not, is predicted on the plausible assumption that a salient context is only available for (21).

To summarize the empirical picture presented in this section, nonatomic distributivity is readily available at the level of the verb (lexical level), but at the level of the verb phrase (phrasal level) it only occurs when context supplies a pragmatically salient cover. Atomic distributivity is available both at the lexical level and at the phrasal level. Summarizing the insights of the previous literature, I assume that this pattern is explained as follows (see Tables 8.1 and 8.2). The lexical cumulativity assumption, encoded in meaning postulates like (19), accounts for the availability of atomic and nonatomic distributivity at the lexical level. Link's atomic D operator is always available at the level of the verb phrase, except in some cases such as *share a pizza*, where its application would lead to nonsensical interpretations. Schwarzschild's cover-based Part operator is also available at the level of the verb phrase, but it is only available if context supplies a salient cover. When this cover contains only one atomic individual in every cell, Schwarzschild's Part operator behaves like Link's D operator.

Even though the semantic effects of D can be subsumed under the workings of Part, I postulate two covert distributivity operators, D and Part, in the grammar. This might seem redundant, and I keep them apart mainly for reasons of symmetry with the two

TABLE 8.1. Distributivity in atomic domains: empirical generalizations

	Lexical (verb level)	Phrasal (verb-phrase level)
atomic	available	available
nonatomic	available	only with context

TABLE 8.2. Distributivity in atomic domains: explanations

	Lexical (verb level)	Phrasal (verb-phrase level)
atomic	lexical cumulativity	atomic operator
nonatomic	lexical cumulativity	cover-based operator

kinds of overt distributive items that are the topic of Chapter 9. For present purposes, it would be equally possible to assume that there is only one operator, namely Part, and that the predicate *Atom* is a salient cover in every situation. Such a view would amount to the following idea: In an atomic domain, the atomic level always provides a salient cover in every context, and this explains the strong preference that speakers have for atomic-level distributivity (see also Rothstein 2010). When the granularity parameter of Part is set to atoms, it behaves equivalently to D. In Chapter 9, I argue that D and Part are lexicalized as adverbial and adnominal distributivity operators in individual languages. This assumption allows us to capture the distinction between English *each* and its German relative *jeweils*. I suggest that the former corresponds to D and the latter corresponds to Part, and I use this assumption to account for the fact that *jeweils* and its relatives across languages have a wider range of readings than *each* and its relatives do.

The search for clear cases of nonatomic distributivity has been going on since at least Link (1987b). Nonatomic lexical distributivity can be observed in the “musicals” example (15), as Lasersohn showed. It also has been argued to occur in examples like the following:

- (27) a. All competing companies have common interests. (Link 1987b)
 b. Five thousand people gathered near Amsterdam. (van der Does 1993)

In example (27a), the predicate *have common interests* can be applied distributively (i.e. it describes several instances of having common interests) to nonatomic entities, because it does not make sense to say of a single company that it has common interests with itself. In example (27b), the predicate *gather near Amsterdam* can be applied distributively (i.e. it describes several gatherings) to nonatomic entities, because a single person cannot gather.

On the other hand, examples that involve *phrasal* nonatomic distributivity, such as the shoe example (21) and the magic square example (22), are harder to come by. I believe that one of the reasons why it has been so hard to identify clear cases of phrasal nonatomic distributivity is the focus in the literature on predicates that apply in count domains. On the standard assumption that the denotations of count nouns are taken from an atomic domain, phrasal distributivity over atoms is naturally expected to be more salient than nonatomic distributivity in almost all contexts and will obscure the presence of nonatomic readings.

The view that atomic granularity is more salient than nonatomic granularity is already defended in Schwarzschild (1996). It can be understood in terms of the principle of Interpretive Economy proposed in Kennedy (2007), which can in turn be derived from first principles in an evolutionary game-theoretic setting (Potts 2008). The central idea is that whenever possible, speakers will converge on certain focal points because this maximizes successful communication. Interpretive Economy was originally proposed to explain why speakers converge on interpreting scalar items

like *tall* and *full* as referring to endpoints of a scale whenever such endpoints exist, and resort to context-dependent values only when this is not the case. In count domains, the scale that results from mapping singular and plural individuals to their cardinalities has an endpoint at 1, the cardinality of singular (atomic) individuals. Interpretive Economy suggests that speakers who use a covert distributivity operator and who need to agree on how to interpret its granularity parameter converge on atomicity as a focal point, except in contexts where another granularity value is salient. This suggests that by looking at noncount domains, we can remove atomic granularity as a potential focal point, so any phrasal distributivity effects we find must be cases of nonatomic distributivity. A reviewer offers the following example from the mass domain as a case in point:

- (28) At the garden party, they sell milk, lemonade, and beer. Milk costs one dollar, lemonade costs two dollars, and beer costs four dollars.

Here, the predicates *cost one dollar* etc. are distributed to a level that is made pragmatically salient by the context, namely the units in which the beverages are sold—presumably glasses, bottles, or cans.

Another nonatomic domain is time. In Section 8.6, I will identify cases of nonatomic phrasal distributivity involving time. I will look at this domain through the lens of *for*-adverbials, focusing on their scopal behavior with respect to verb phrases that contain an overt quantifier. In a nonatomic domain, there are necessarily no atomic covers, so the first rows of Tables 8.1 and 8.2 are irrelevant here. I will argue that the second rows of these tables are mirrored precisely in the temporal domain, as shown in Tables 8.3 and 8.4.

As described in Chapter 4, the fundamental assumption of strata theory is that distributivity always involves a dimension and a granularity parameter. That is, the

TABLE 8.3. Distributivity in nonatomic domains: empirical generalizations

	Lexical (verb level)	Phrasal (verb-phrase level)
atomic	n/a	n/a
nonatomic	available	only with context

TABLE 8.4. Distributivity in nonatomic domains: explanations

	Lexical (verb level)	Phrasal (verb-phrase level)
atomic	n/a	n/a
nonatomic	lexical cumulativity	cover-based operator

dimension parameter of the distributivity operator involved can be instantiated to τ (runtime) and, in that case, its granularity parameter is dependent on an anaphorically salient level of granularity. I have suggested that the distributivity operators contain these two parameters. The granularity parameter can be understood as Schwarzschild's cover, but the dimension parameter does not yet figure in the implementations we have seen so far. Therefore, it is necessary at this point to provide a formal implementation of Link's and Schwarzschild's distributivity operators that supports the notion of dimension and granularity parameters. Section 8.4 provides this implementation.

8.4 Reformulating the D and Part operators

In their original formulation, the distributivity operators I have discussed so far only distribute over the subject, and they return a truth value once they have combined with it. But as we have seen above, examples like (7)—repeated here as (29)—show that these operators must be able to target different argument positions, and that they can combine with more than one constituent:

- (29) a. The first-year students [D [took an exam]]. *Target: agent*
 b. John [D [gave a pumpkin pie]] to two girls. *Target: goal*

Lasersohn (1998b) discusses how to generalize these operators in a way that allows them to distribute over different argument positions. But if the previous discussion is on the right track, distributivity does not have to target overt argument positions. It can also target salient temporal intervals that are not denoted by any argument in the sentence. This cannot easily be accommodated on the traditional Montagovian view, unless we add a silent temporal argument to each verb. A simpler and more uniform picture is provided by Neo-Davidsonian event semantics, on which thematic roles like *agent* and related functions like *runtime* are reified. On this view, it makes sense to think of the D and Part operators as being parametrized on these functions.

Once we move to event semantics, the distributivity operators developed and motivated in the last two sections have to be adjusted for a number of reasons. First, Link's and Schwarzschild's formulations of distributivity operators assume that verb phrases denote sets of individuals, while event semantics typically assumes that verb phrases denote sets of events (see Section 2.10). Second, like other kinds of quantification in event semantics, distributivity over individuals requires a variable that ranges over subevents of some sum event, and this needs to be built into the operators (Taylor 1985; Schein 1993). Third, as we will see, there is a technical flaw with the main existing proposal of how to represent the D operator in event semantics, namely Lasersohn (1998b). Specifically, Lasersohn's operator fails to prevent the sum event from containing extraneous material in addition to the subevents over which the D operator ranges ("leakage," in the terms of Bayer (1997)). This problem also applies to the Part operator.

The star operator, defined in Section 2.3.4, will play a central role in the development to follow. Distributivity is usually seen as involving universal quantification, and as we will see, the original definitions of Link's (12) and Schwarzschild's (24) operators contain quantifiers. My distributivity operators will be built around the star operator instead. This is not as big a shift as it might seem. As discussed in Section 2.3.4, there is a close relation between the star operator and universal quantification.

Recall from Section 2.7.2 that not all verb phrases have cumulative reference. For example, the sum of two events in the denotation of the verb phrase *be paid exactly \$7,000* will not, as a general rule, again be in its denotation, because it will usually involve \$14,000 rather than \$7,000. Likewise, the sum of two events in the denotation of the verb phrase *find a flea* will only end up in its denotation if the two events happen to involve the same flea. This fact is important because of the distinction between lexical and phrasal distributivity. If all verb phrases were cumulative, then under the present assumptions there would be no way to explain why nonatomic distributivity is not readily available at the verb phrase level, as discussed in Section 8.3.

Let me now reformulate the D and Part operators in a way that meets the requirements mentioned earlier. Section 8.5 compares it with the proposal by Lasnik (1998b). The main idea is that this operator shifts any predicate (typically a verb phrase) into a predicate that satisfies a condition that is analogous to the one captured in the meaning postulates discussed in Section 4.5.1.

The key idea that underlies my reformulation is that the two operators may be seen as special cases of stratified reference. Since stratified reference tells us what it means for a predicate to be distributive, the purpose of the D operator can be understood as shifting arbitrary event predicates to an interpretation that has stratified reference. The definition of stratified reference is repeated here from Section 4.6, but with a change: I have λ -abstracted over P and x .

(30) **Definition: Stratified reference**

Let d (a *dimension*) be any function from entities of type α to entities of type β , and let g (a *granularity level*) be any predicate of entities of type β . Let P range over predicates of type $\langle \alpha, t \rangle$ where α is either e or v , and let x range over entities of type α . Then:

$$SR_{d,g} \stackrel{\text{def}}{=} \lambda P \lambda x. x \in * \lambda y \left(\begin{array}{c} P(y) \wedge \\ g(d(y)) \end{array} \right)$$

(P stratifies x along dimension d with granularity g iff x consists of one or more parts in P that are each mapped by d to something in g .)

From this definition, we obtain the D_θ operator in (31) by setting α to v , β to e , d to θ , and g to *Atom*, and by coindexing the operator with θ :

(31) **Definition: Event-based D operator**

$$[D_\theta] \stackrel{\text{def}}{=} \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$$

This operator applies to an event predicate V , such as a verb phrase. It returns another event predicate, one which holds of any event e as long as it consists of one or more events that are in V and which are each mapped by the function θ to an atom. I assume that θ is a free variable that is resolved to a thematic role—formally, a function that maps events to their agents, themes, and so on. As discussed in Section 2.5.1, I adopt thematic uniqueness, and I model thematic roles as partial functions. This means that they can be treated as being the same kinds of things as functions that map events to their spatial and temporal locations. This will be useful once we extend the D operator to nonatomic domains like time.

Examples (32) and (33) illustrate how the D operator in (31) works. Sentence (32) gives a baseline, a scopeless reading that does not use the D operator. Sentence (33) shows the D operator in action to model a distributive reading.

(32) The boys saw a monkey.

$$\exists e[*agent(e) = \bigoplus boy \wedge *see(e) \wedge monkey(theme(e))]$$

(There is a potentially plural seeing event whose agents sum up to the boys, and whose theme is one monkey. That is, only one monkey is seen.)

(33) The boys $[D_{agent} [saw a monkey]]$.

$$\exists e. *agent(e) = \bigoplus boy \wedge$$

$$e \in * \lambda e' \left(\begin{array}{l} *see(e') \wedge \\ monkey(theme(e')) \wedge Atom(agent(e')) \end{array} \right)$$

(There is an event whose agents sum up to the boys, and this event consists of seeing events for each of which the agent is an atom and the theme is a monkey.)

The star operator $* \lambda e'$ is introduced through the D operator and takes scope over the predicate *monkey* introduced by the theme. The representation (33) does not state explicitly that each boy sees a monkey, so it might not be clear that it is an adequate way to capture what the sentence means. This is where the background assumptions introduced in Chapter 2 come into play. The representation (33) explicitly states that the monkey-themed events e' have atoms as agents. The fact that these atoms are boys is entailed by the background assumption that the entities in the denotation of singular count nouns are atoms (Section 2.6.1), together with the background assumption that thematic roles are cumulative (Section 2.5.1).

Because the D operator in (31) carries an index, we can capture through a simple change in coindexation the kinds of configurations that have otherwise been taken to require movement or type shifting. For example, the reading of (29b)—repeated below—in which each of the two girls gets a pumpkin pie from John can be straightforwardly accounted for.

(34) John $[D_{goal} [gave a pumpkin pie]]$ to two girls. = (29b)

(35) $[[give a pumpkin pie]] = \lambda e[*give(e) \wedge pumpkin-pie(theme(e))]$

$$(36) \llbracket \text{to two girls} \rrbracket = \lambda e [\text{two-girls}(*\text{goal}(e))]$$

$$(37) \llbracket D_{\text{goal}} \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\text{goal}(e')) \end{array} \right)$$

$$(38) \llbracket D_{\text{goal}} \text{ give a pumpkin pie} \rrbracket = (37)((35)) \\ = \lambda e. e \in * \lambda e' \left(\begin{array}{c} * \text{give}(e') \wedge \text{pumpkin-pie}(\text{theme}(e')) \wedge \\ \text{Atom}(\text{goal}(e')) \end{array} \right)$$

$$(39) \llbracket [D_{\text{goal}} \text{ give a pumpkin pie}] \text{ to two girls} \rrbracket = (36) \cap (38) \\ = \lambda e. \text{two-girls}(*\text{goal}(e)) \wedge \\ e \in * \lambda e' \left(\begin{array}{c} * \text{give}(e') \wedge \text{pumpkin-pie}(\text{theme}(e')) \wedge \\ \text{Atom}(\text{goal}(e')) \end{array} \right)$$

$$(40) \llbracket (34) \rrbracket = \exists e. \text{agent}(e) = \text{john} \wedge \text{two-girls}(*\text{goal}(e)) \wedge \\ e \in * \lambda e' \left(\begin{array}{c} * \text{give}(e') \wedge \text{pumpkin-pie}(\text{theme}(e')) \wedge \\ \text{Atom}(\text{goal}(e')) \end{array} \right)$$

(There is an event whose agent is John, whose goal is two girls, and which consists of parts each of which is a giving event that has an atomic goal and a pumpkin pie as its theme.)

Turning now to the event-based reformulation of Schwarzschild's Part operator, this one can be seen as a generalization of Link's D operator. Instead of specifying the granularity parameter to be atomic, we leave it free. Accordingly, we obtain the reformulation by replacing *Atom* in (31) with a free variable *C*, which I will assume is anaphoric on an antecedent that can be provided by the context. This minimal change reflects the close connection between D and Part.

(41) **Definition: Event-based Part operator**

$$\llbracket \text{Part}_{\theta, C} \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right)$$

What this says is that Part takes an event predicate *V* and returns a predicate that holds of any event *e* which can be divided into events that are in *V* and whose *θ*s satisfy the contextually salient predicate *C*. Whenever *θ* is a sum homomorphism, this will entail that the *θ*s of these events sum up to the *θ* of *e*.

Definition (41) entails that *C* is a set whose sum is *θ(e)*. In this way, the notion of cover emerges naturally and does not need to be separately defined. This may be seen as a conceptual advantage over Schwarzschild (1996) (see Section 5.4).

The example in (42) show how my reformulation of Schwarzschild's Part operator works. The example is repeated from the shoe sentence (2). I assume for concreteness that the shoes play the theme role in this sentence. I assume that *C* is resolved here to a contextual predicate that I call *pair* and that applies to a sum of two shoes just in case they are one of the pairs of shoes on display.

(42) The shoes [Part_{theme,pair} [cost fifty dollars]].

$$\exists e.^* \text{theme}(e) = \bigoplus \text{shoe} \wedge \text{Part}_{\text{theme,pair}}(\llbracket \text{cost fifty dollars} \rrbracket)(e)$$

$$\Leftrightarrow \exists e.^* \text{theme}(e) = \bigoplus \text{shoe} \wedge$$

$$e \in {}^* \lambda e' \left(\begin{array}{l} e' \in \llbracket \text{cost fifty dollars} \rrbracket \wedge \\ \text{pair}(\text{theme}(e')) \end{array} \right)$$

(There is a plural event whose themes sum up to the shoes and which consists of costing-fifty-dollars events with pairs as themes.)

The event-based reformulation of the D and Part operators allows us to think of distributivity as a parametrized operator. The D operator only has one parameter, θ , which specifies the thematic role over which it distributes. The Part operator has an additional granularity parameter, C , which corresponds to Schwarzschild's cover variable and which specifies the things over which it distributes. These two parameters are at the core of strata theory, as discussed in Section 8.1 and in Champollion (2010b). The dimension parameter of strata theory corresponds to the θ parameter on our operators, and it indicates the domain that contains the entities over which the operator distributes. The granularity parameter of strata theory corresponds to the C parameter of the Part operator, and it indicates the size of the entities over which the operator distributes. I assume that the setting "granularity=atom" blocks the setting "dimension=time" because time is continuous and noncount (see Section 2.4.4). That is, either there are no atoms to distribute over, or they are not accessible to natural language semantics (von Stechow 2009). In Chapter 9, I rely on this interaction in order to explain certain typological facts involving distance-distributive items.

This similarity to stratified reference is not the only reason why I adopt the definitions in (31) and (41) over alternative formulations. Another reason is that these definitions provide access to the sum of the events over which the relevant predicate is distributed. This sum event will be of central importance throughout the rest of this chapter and Chapter 9. It is absent not only from traditional definitions of these operators (because they do not invoke events), but also from previous proposals on how to reformulate distributivity operators in event semantics. The next section reviews one of these proposals.

8.5 Previous work on event-based distributivity operators

Reformulating distributivity operators in event semantics is one of the topics of Lasnik (1998b). That paper shows a way of generalizing Link's and other distributivity operators so that they apply to other positions than the subject position both in eventless and in event-based frameworks. The following entry is a special case among these different combinations, namely the VP-level variant of an event-based version of Link's D operator.

(43) **Distributivity operator over events (Lasersohn)**

$$[[D]] = \lambda P_{\langle e, vt \rangle} \lambda x \lambda e \forall y [y \leq_{\text{Atom}} x \rightarrow \exists e' [e' \leq e \wedge P(y)(e')]]$$

Lasersohn's operator applies to a predicate of type $\langle e, vt \rangle$. It is based on the assumption that a verb phrase like *smile* that is about to combine with it is represented as something like $\lambda x \lambda e [\text{smile}(e) \wedge \text{agent}(e) = x]$. The entry in (43) does not represent Lasersohn's entire proposal, but it is the part that is most closely related to mine. A very similar operator is proposed in LaTerza (2014a). Both authors credit the basic idea behind their event-based operators to Schein (1993). In the following, for concreteness I only talk about Lasersohn's proposal, but my observations apply equally to LaTerza's operator.

One difference between Lasersohn's proposal in (43) and my proposal in (31) consists in the way the operators access the plural entity over which they distribute. The D operator in (43) accesses that plural entity by combining with a predicate of type $\langle e, vt \rangle$. In contrast to Lasersohn, I do not rely on the implicit assumption that the D operator is adjacent to a constituent denoting that entity. My operator in (31) is parametrized for the relevant thematic role. This thematic role can be supplied by coindexation with an appropriate θ role head (I come back to this in Section 9.2). Maintaining the adjacency assumption would make it harder to build on the D operators to account for the phenomenon of distance distributivity, as I do in Chapter 9.

Another, more substantial difference consists in the way in which the operators access the events over which they distribute. My operator in (31) uses algebraic closure over events with atomic agents, while Lasersohn's operator uses universal quantification over individuals. The difference between the formulations is apparent in the different representations that result from inserting a D operator into the sentence *The children took a nap* before existential closure applies. (The subformula $y \leq_{\text{Atom}} \bigoplus \text{child}$ is equivalent to $\text{child}(y)$ on the assumption that children are mereological atoms.)

(44) a. **Lasersohn's representation:**

$$\lambda e \forall y [y \leq_{\text{Atom}} \bigoplus \text{child} \rightarrow \exists e' [e' \leq e \wedge * \text{take}(e') \wedge \text{agent}(e') = y \wedge \text{nap}(\text{theme}(e'))]]$$

b. **My representation:**

$$\lambda e [* \text{agent}(e) = \bigoplus \text{child} \wedge e \in * \lambda e' [* \text{take}(e') \wedge \text{Atom}(\text{agent}(e')) \wedge \text{nap}(\text{theme}(e'))]]$$

Lasersohn's representation (44a) applies to all events that contain a napping subevent (i.e. a "taking" subevent whose theme is a nap) for each child, even if they also contain other subevents. My representation applies to all events that contain a napping subevent for each child and nothing else. Thus, Lasersohn's solution suffers from leakage. It does not give special status to the sum of all of the subevents over

which the D operator distributes. Instead, it applies not only to that sum but also to any event that contains that sum. Through leakage, Lasersohn's operator makes any predicate it modifies persistent. That is, whenever the modified predicate applies to an event e , it also applies to any event of which e is a part. Eckardt (1998: ch. 4) argues convincingly that the persistency assumption is problematic in event semantics, contra Lasersohn (1992, 1998b). I spell out one facet of this problem shortly. By contrast, my representation (44b) only applies to the sum itself. This is due to the way the star operator works. To expand on the geometrical example given in Section 2.3.4, if x is a square and P is the property of being a triangle, then x satisfies $*P$ because x can be divided into triangles *without leaving anything out*. If x is a circle instead of a square, it does not satisfy $*P$. Even though every circle contains infinitely many triangles, the curvature of the circle makes it impossible to divide it into triangles without leaving anything out (setting aside sums of uncountably infinite triangles).

Leakage causes a problem in connection with subjects and other arguments and modifiers that take scope over the distributivity operator. This will be relevant in Section 8.6, where I will argue that the distributivity operator can also occur in the scope of *for*-adverbials. The problem can be illustrated with nondistributive adverbials such as *in 30 minutes* and *from 2pm to 4pm* (see Section 2.4.3 and Eckardt 1998: chs 4 and 5) or *unharmoniously* (Schein 1993: ch. 1):

- (45) a. In 30 minutes, Alma put each ball into a box.
 b. From 2pm to 4pm, Bertha took a nap and watered the tulips.
 c. Unharmoniously, every organ student sustained a note on the Wurlitzer.

The predicates in these examples hold of an event even if they do not hold of its parts. For example, (45c) is true if the ensemble event was unharmonious even if the same cannot be said of any one student's note. We can now illustrate the problem caused by leakage. Let L stand for (44a) and let M stand for my event predicate (44b). Imagine a group of children in a preschool napping from 2pm to 3pm and then playing from 3pm to 4pm. For each child there is a napping event followed by a playing event. Call the sum of all the napping events e_{nap} . By lexical cumulativity, this sum itself counts as a napping event, and its agent is the children; hence it satisfies both Lasersohn's predicate L and my predicate M . Call the sum of all playing events e_{play} , a playing event whose agent is again the children. Let $e_{nap} \oplus e_{play}$ be the sum of e_{nap} and e_{play} . This does not satisfy M because it contains extraneous material: it is not a napping event but the sum of a napping and of a playing event, and the two do not overlap. But $e_{nap} \oplus e_{play}$ satisfies L , because by virtue of containing e_{nap} , it contains a napping event for every child. By assumption, e_{nap} does not take place from 2pm to 4pm, but $e_{nap} \oplus e_{play}$ does. Sentence (46) is false in this scenario. If the D operator is applied to *nap*, then on Lasersohn's account, this sentence is represented as (46a), while on my account it is represented as (46b).

- (46) From 2pm to 4pm, the children took a nap.
 a. $\exists e[e \in \llbracket \text{from 2pm to 4pm} \rrbracket \wedge L(e)]$
 b. $\exists e[e \in \llbracket \text{from 2pm to 4pm} \rrbracket \wedge M(e)]$

Now $e_{\text{nap} \oplus \text{play}}$ satisfies both L (by leakage) and the predicate *from 2pm to 4pm* (by assumption). Therefore, Lasersohn's D operator wrongly predicts that (46) is judged true, namely in virtue of $e_{\text{nap} \oplus \text{play}}$.

The reader may wonder whether the leakage problem could be avoided by combining Lasersohn's operator with minimization or some related operation. This will not work, because events are wholly but not necessarily minimally relevant to the truth of the propositions they verify (see Fine 2012). Let $\min(V)$ denote the set of all entities e such that V applies to e but not to any of the proper parts of e , and consider the following amendment to Lasersohn's operator (43):

- (47) $\lambda P_{\langle e, vt \rangle} \lambda x \lambda e. e \in \min(\lambda e' \forall y[y \leq_{\text{Atom}} x \rightarrow \exists e''[e'' \leq e' \wedge P(y)(e'')]])$

When this version is inserted into *The children took a nap*, the resulting representation before existential closure applies to all events that are minimal with respect to the property of containing a napping subevent for each child. In the example above, $e_{\text{nap} \oplus \text{play}}$ will not be such an event, because it is not minimal with respect to that property. This is as desired.

But in the context of event semantics, the minimality operation has unintended consequences. For one thing, there may not always be any minimal events. Take a predicate with the subinterval property, such as *sit*. If we assume that *sit* is distributive on its agent position and that time is nonatomic, applying (47) to it returns the empty set because there are no minimal sitting events. A possible reaction is to replace the minimality condition in (47) by an exemplification condition (Kratzer 2016). An entity e exemplifies a predicate V just in case V holds of e and either of none of the proper parts of e (as in the case of minimality) or of all of them. This avoids the problem just mentioned. However, both minimality and exemplification fail to give us access to events of the required length. Even if minimal napping events exist, their runtime will certainly be shorter than one hour. This means that in the scenario above where each child takes a nap from 2pm to 3pm, *The children took a nap from 2pm to 3pm* is predicted false, contrary to fact. Even more sophisticated notions of minimality have also turned out to be problematic (Eckardt 1998: ch. 4). Rather than pursuing this route further, I will adopt the operator in (31). This operator not only avoids the problems just mentioned. It is also more simple and concise than (47) plus exemplification, and as we have seen, it makes the parallel with strata theory clear.

As this section has shown, to avoid leakage we need to reformulate the distributivity operator in a way that makes sure that no extraneous material can find its way into the sum event. This ensures that the output of the operator is the right kind of predicate to be passed on to the next argument or adjunct of the verb. In the previous examples,

this was the event modifier *from 2pm to 4pm*. In the next section, I will exploit this property of the distributivity operator in another way: it will apply to a verb phrase as before, but its output will serve as the input to a *for*-adverbial, which in turn will pass it on to the subject of the sentence after making sure that the output of the distributivity operator is atelic.

8.6 The scopal behavior of *for*-adverbials

We have seen in Sections 8.2 and 8.3 that covarying interpretations of indefinites and numerals are the signature property of phrasal distributivity. In this respect, covert phrasal distributivity is of course similar to overt universal quantification:

- (48) a. The girls are wearing a black dress.
 b. Every girl is wearing a black dress.
 c. Every day, Mary wore a black dress.

As we will see in this section, *for*-adverbials are often analyzed as involving universal quantification over temporal intervals. From this point of view, one would expect them to act similarly to overt universal quantifiers like *every girl* or *every day*. Yet it has often been observed that they do not behave the same way as overt universal quantifiers do (e.g. Carlson 1977; Zucchi & White 2001; van Geenhoven 2004; Kratzer 2007). I will argue that phrasal distributivity is not triggered by *for*-adverbials, but by intervening modifiers like *every day* and by distributivity operators that arise in contexts where a temporal partition is salient.

Normally, *for*-adverbials cannot give rise to covariation in singular indefinites. The sentences in (49), adapted from Kratzer (2007), are supposed to be understood as uttered out of the blue, without any special context. I write $\exists > \forall$ for an interpretation in which the indefinite involves reference to a single entity and $\forall > \exists$ for an interpretation in which it involves reference to multiple entities.

- | | |
|---|--|
| (49) a. I dialed a wrong phone number for five minutes. | $\exists > \forall$; $*\forall > \exists$ |
| b. John pushed a cart for an hour. | $\exists > \forall$; $*\forall > \exists$ |
| c. She bounced a ball for 20 minutes. | $\exists > \forall$; $*\forall > \exists$ |
| d. He kicked a wall for a couple of hours. | $\exists > \forall$; $*\forall > \exists$ |
| e. She opened and closed a drawer for half an hour. | $\exists > \forall$; $*\forall > \exists$ |
| f. I petted a rabbit for two hours. | $\exists > \forall$; $*\forall > \exists$ |

It would be plausible for (49a) to have an interpretation like *Over and over again over the course of five minutes, I dialed a different wrong phone number*, and similarly for the other examples. But this kind of interpretation is systematically absent from such examples when they are uttered out of the blue. Even in cases where the wide-scope interpretation of the indefinite is pragmatically odd, it is still the only one available. Example (50a) is repeated from (4a); (50b) and (50c) are from Deo & Piñango (2011).

- (50) a. ??John found a flea for a month. $??\exists > \forall; *\forall > \exists$
 b. ??John noticed a discrepancy for a week. $??\exists > \forall; *\forall > \exists$
 c. ??John discovered a new proof for a week. $??\exists > \forall; *\forall > \exists$

The same behavior can be observed if we replace the singular indefinite with certain other types of quantifiers, such as plural indefinites. For example, the only available interpretation of (51) is the one in which the plural indefinite *thirty zebras* involves reference to a single set of thirty zebras.

- (51) John saw thirty zebras for three hours.

Even in German, a language in which inverse scope is normally not available, indefinites in the scope of *for*-adverbials cannot covary. Each of the following two sentences is a possible translation of (49a) and must involve reference to a single phone number (Kratzer 2007).

- (52) Ich hab' fünf Minuten lang eine falsche Telefonnummer gewählt.
 I have five minutes long a wrong telephone.number dialed
 (53) Ich hab' eine falsche Telefonnummer fünf Minuten lang gewählt.
 I have a wrong telephone.number five minutes long dialed

The behavior just described does not hold across the board for all types of noun phrases. Bare plurals and mass nouns do not have to take distributive wide scope over *for*-adverbials (Carlson 1977; Verkuyl 1972; Dowty 1979). This can be seen in the sentences in (54) and (55), taken from Dowty (1979):

- (54) a. John found fleas on his dog for a month.
 b. John discovered crabgrass in his yard for six weeks.
 (55) a. Tourists discovered that quaint little village for years.
 b. Water leaked through John's ceiling for six months.

VP-level and sentential predicates with bare noun phrases are generally compatible with *for*-adverbials and do not give rise to the phenomena described above. For example, (54a) is compatible with the plausible interpretation in which John finds different fleas on his dog over the course of a month and finds each of them only once.

Even singular indefinites in the scope of *for*-adverbials can involve reference to multiple individuals when a temporal universal quantifier intervenes, as in examples (5) and (6b), repeated here as (56a) and (56b).

- (56) a. John found a flea on his dog every day for a month.
 b. Jim hit a golf ball into the lake every five minutes for an hour.

In these examples, the indefinite can covary with the universal quantifier. As we have seen, (56a) is true in a situation where John found a different flea on his dog every day,

and (56b) is compatible with Jim hitting a different golf ball into the lake every five minutes. Other interveners allow indefinites under *for*-adverbials to involve reference to multiple entities. These include pluractional adverbials like *day after day* (see Beck & von Stechow 2007) and context-dependent temporal definites such as *at breakfast* (Deo & Piñango 2011):

- (57) a. John found a flea on his dog day after day for a month.
b. Jane ate an egg at breakfast for a month.

Covariation is also possible when a salient level of granularity can be inferred from context (see Ferreira 2005: 130). To repeat the example from Section 8.1, in a context where the daily pill intake of patients is salient, such as a hospital, sentence (58) is licit despite the fact that it does not require any pill to be taken more than once.

(58) **Context: discussing daily pill intake**

The patient took two pills for a month and then went back to one pill.

Example (59), from Deo & Piñango (2011), shows the same point. It is understood as involving reference to several snowmen.

- (59) We built a huge snowman in our front yard for several years.

This is presumably because world knowledge makes the cycle of seasons salient here, suggesting that they built a different snowman every winter.

Finally, example (60) is adapted from Landman & Rothstein (2009) and is to be understood in a context where the bicycle is designed to carry around three children at a time, and was used over a period of twenty years to carry different sets of three children around.

- (60) This bicycle carried three children around Amsterdam for twenty years.

In the following, τ stands for the *runtime* function that maps each event to the location in time at which it occurs. As discussed in Section 2.5.2, I assume that τ is a sum homomorphism, which means that runtimes can be discontinuous. Take for example an event e_1 whose runtime is the interval from 12.30pm to 1pm, and an event e_2 whose runtime is the interval from 5pm to 6.15pm. The sum of these two events, $e_1 \oplus e_2$, will have a discontinuous runtime, namely the sum of the interval from 12.30pm to 1pm and the interval from 5pm to 6.15pm. Lexical cumulativity causes most verbs to involve reference to many discontinuous events; for example, *find* applies to sums of finding events that may have arbitrarily long gaps between them.

Theories of *for*-adverbials can be classified into two groups, which I will review here schematically as Theory A and Theory B. Theory A predicts that all indefinites in the scope of *for*-adverbials should covary, while Theory B predicts that none of them should. Neither of these theories turns out to account for the facts by itself.

I will propose to account for the limited covariation by adding a distributivity operator to Theory B. Roughly these two kinds of theories (but without the distributivity operator) are also discussed by Zucchi & White (2001). In my discussion of the two classes of theories I will use two shorthands, *regular* and *atelic*, to hide some complexities that do not affect the main point of comparison. The shorthand *regular* was explained in Section 3.3. The shorthand *atelic* stands for the constraint that makes *for*-adverbials rule out telic predicates. I use it to remain neutral between other implementations of this constraint and my own. I have argued in Chapters 5 and 6 that it can be formulated in terms of stratified reference, with a dimension parameter and a granularity parameter.

On Theory A, the meaning of a *for*-adverbial can be represented as something like (61):

$$(61) \quad \llbracket \text{for an hour} \rrbracket (\text{Theory A}) \\ = \lambda V \exists t [\text{hours}(t) = 1 \wedge \text{regular}(t) \wedge \\ \forall t' [t' \text{ is a very short part of } t \rightarrow \exists e [V(e) \wedge \tau(e) = t']]]$$

Theory A says that a *for*-adverbial is represented as a universal quantifier over very short subintervals of a regular interval, a bit as if it was the temporal counterpart of *every*. By contrast, Theory B says that a *for*-adverbial is represented in a way that passes the denotation of the verb phrase up unchanged as long as it is both regular and atelic. On Theory B, the meaning of a *for*-adverbial can be represented as in (62) (recall that I write $\lambda e : \phi. \psi$ for the partial function that is defined whenever ϕ holds, and that maps e to ψ whenever it is defined):

$$(62) \quad \llbracket \text{for an hour} \rrbracket (\text{Theory B}) \\ = \lambda V \lambda e : \text{regular}(\tau(e)) \wedge \text{atelic}(V). V(e) \wedge \text{hours}(\tau(e)) = 1$$

Theory A is somewhat similar to the influential theory of *for*-adverbials in Dowty (1979) discussed in Section 5.3.1. However, there is an important difference, because Dowty analyzes *for*-adverbials as universal quantifiers over subintervals rather than over instants. These do not have to be proper subintervals, so the predicate is required to hold at the entire interval described by the *for*-adverbial, in addition to all of its subintervals. A similar theory, in which *for*-adverbials quantify over “relevant” subintervals (whatever that may mean), is found in Moltmann (1991) and discussed in Section 5.3.3. Theory B is found in various forms in Krifka (1989a, 1989b, 1998). It is explicitly defended against Theory A in Kratzer (2007).

The behavior of indefinites in examples like (49) and (50), some of which are repeated in (63), is surprising on Theory A, where a *for*-adverbial *for an hour* is interpreted as *at each very short part of the relevant interval*:

- (63) a. I dialed a wrong phone number for five minutes. = (49a)
 b. ??John found a flea for a month. = (50a)

On Theory A, in order to account for this behavior the indefinite would have to be interpreted with wide scope over the *for*-adverbial. In (63b), the resulting interpretation is pragmatically odd because it is unusual to find the same flea repeatedly. The narrow-scope interpretation would be much more plausible, but it is not available out of the blue. One might try to account for these facts by stipulating obligatory quantifier raising of the indefinite above the *for*-adverbial, as suggested by Krifka (1998). But this will not work in German, a surface-scope language whose *for*-adverbials behave as in English, as we have seen in (52) and (53) (Kratzer 2007). Moreover, if *for*-adverbials forced all indefinites in their syntactic scope to move and take semantic scope above them, then *every day* should not be able to prevent this from happening in this case, contrary to what we have seen in (56a) (Zucchi & White 2001). Thus, obligatory quantifier raising is not an option.

Neither Theory A nor Theory B is able to account for the limited ability of indefinites to covary in the scope of *for*-adverbials, at least not without further modifications. In a nutshell, Theory A predicts that they should always covary and Theory B predicts that they should never covary. On Theory A, the scopal behavior of *for*-adverbials is surprising when compared with the familiar scopal behavior of the universal quantifier *every*. In contrast to *every*, which can take scope anywhere in its clause, *for*-adverbials always seem to take narrow semantic scope with respect to singular indefinites in their syntactic scope, except in cases like (56a) through (60). Thus the ability of *for*-adverbials to give rise to quantifier scope ambiguities is much more limited than we would expect. Theory B, on the other hand, is not a good fit either, because by itself it predicts no scope ambiguities at all.

Clearly, neither of the two classes of theories can model the covariation of indefinites while taking the role of context into account. The first part of this chapter has already introduced the tool we need to account for contextually limited covariation: a nonatomic distributivity operator. I will assume that this operator can intervene between the verb phrase and the *for*-adverbial in the same way as overt adverbs like *every day/hour/month*, and with a similar effect on the behavior of indefinites in their scope. In a nonatomic domain like time, atomic distributivity is not an option; so when the dimension parameter of the distributivity operator involved is instantiated to τ (runtime), its granularity parameter cannot be set to *Atom*. I assume that nonatomic levels of granularity such as *day* and *month* are only available in special contexts that make them salient, just as in the case of Schwarzschild's nonatomic Part operator. On this view, then, indefinites that seem to covary with a *for*-adverbial actually covary with a covert Part operator.

I will now implement my proposal by extending Theory B with a temporal version of the Part operator, which is related to the notion of a contextually determined partition that originated independently in Moltmann (1991) and in Deo (2009). These accounts place the anaphoricity on context into the *for*-adverbial itself. This approach is also advocated in Deo & Piñango (2011). I have argued elsewhere that it is the

distributivity operator and not the *for*-adverbial that is anaphoric on context (Champollion 2013). As we have seen, the Part operator introduces a contextual variable that is resolved to a cover. I will identify this cover with the granularity parameter *day/hour/month*.

My concrete implementation of Theory B will be based on the translation of *for*-adverbials in Section 4.7, with the stratified-reference presupposition implementing the atelicity requirement (and with regularity turned into a presupposition):

- (64) $\llbracket \text{for an hour} \rrbracket$
 $= \lambda P_{\langle v, t \rangle} \lambda e : \text{regular}(\tau(e)) \wedge \text{SR}_{\tau, \gamma(\tau, e)}(P)(e).$
 $P(e) \wedge \text{hours}(\tau(e)) = 1$

Let us first consider cases where Theory B works without modifications, namely those in which the indefinite does not covary. Given our background assumptions, Theory B immediately predicts that the indefinite in (49b), repeated here as (65), must involve reference to a single cart.

- (65) John pushed a cart for an hour. = (49b)

This prediction is obtained based on the representation in (66) of the denotation of the verb phrase *push a cart*, which assumes lexical cumulativity, as in (61):

- (66) $\llbracket \text{push a cart} \rrbracket = \lambda e [* \text{push}(e) \wedge \text{cart}(\text{theme}(e))]$
 (True of any pushing event or sum of pushing events whose theme is one and the same cart.)

Even though the verbal denotation *push* is pluralized here, the predicate *cart* is not pluralized. That is, the verb phrase only applies to events whose theme is exactly one cart, even if these events may be sums of other events. In connection with the entry (62), it predicts that the entire event over which sentence (65) existentially quantifies must have a single cart as its theme:

- (67) $\llbracket \text{John pushed a cart for an hour} \rrbracket$
 $= \exists e : \text{regular}(\tau(e)) \wedge \text{SR}_{\tau, \gamma(\tau, e)}(\lambda e' [* \text{push}(e') \wedge \text{cart}(\text{theme}(e'))])(e).$
 $[* \text{agent}(e) = j \wedge * \text{push}(e) \wedge \text{cart}(\text{theme}(e)) \wedge \text{hours}(\tau(e)) = 1]$
 (Presupposes that there is a regular event that is stratified by the property of pushing a cart. Asserts that this event is a pushing event whose theme is one cart, whose agent is John, and whose runtime measures one hour.)

Lexical cumulativity accounts for the behavior of achievement verbs like *find* even though these verbs are normally understood to have very short runtimes (Kratzer 2007). For example, a sentence like (68) (repeated from (50a)) is now predicted to entail that there was a finding event *e* which lasted a month and whose theme is a flea.

- (68) $\llbracket \text{John found a flea for a month} \rrbracket$
 $= \exists e : \text{regular}(\tau(e)) \wedge \text{SR}_{\tau, \gamma(\tau, e)}(\lambda e' [* \text{find}(e') \wedge \text{flea}(\text{theme}(e'))])(e).$
 $[* \text{agent}(e) = j \wedge * \text{find}(e) \wedge \text{flea}(\text{theme}(e)) \wedge \text{months}(\tau(e)) = 1]$
 (Presupposes that there is a regular event that is stratified by the property of finding a flea. Asserts that this event is a finding event whose theme is one flea, whose agent is John, and whose runtime measures one month.)

The lexical cumulativity assumption allows this finding event to consist of several individual findings, which may have very short runtimes. It allows phrasal predicates like *find a flea* to involve reference to plural events only to the extent that the verb predicate (*find* in this case) already does so. The object *a flea* is not affected by pluralization and continues to involve reference to a singular flea. This means that sentence (68) requires a single flea to have been found repeatedly over the course of a month.

The formula in (68) does not require John to have been searching uninterruptedly at every moment of the month, since the runtimes of events may be discontinuous. I assume that the function *months* maps a discontinuous interval to the same number as the smallest continuous interval that contains it (see Section 2.5.4).

Without further modifications, Theory B predicts that a singular indefinite in the scope of *for*-adverbials should always involve reference to a single entity, consistent with what we have seen in the examples in (49) but contrary to what we have seen in examples (58), (59), and (60). The only criterion that distinguishes these two groups of examples is the availability of a supporting context. We observed an analogous effect in Section 8.3, when we considered the following examples:

- (69) a. The shoes cost fifty dollars (*i.e. per pair*). = (21)
 b. The suitcases weigh fifty pounds. = (23)

I will now add a distributivity operator to Theory B that is anaphoric on a contextually salient temporal cover. This will allow context to rescue examples like (58), (59), and (60), in the same way as examples like (69a) involve a distributivity operator over shoes which is anaphoric to a salient cover of the collection of shoes in question. Out of the blue, predicates like *find a flea* in examples like (50a) are unable to distribute over days for the same reason that also prevents predicates like *weigh fifty pounds* in examples like (69b) from distributing over pluralities of suitcases. Both would need a Part operator in order to distribute, but that operator is not licensed out of the blue.

The distributivity operator in Section 8.4 was relativized to two parameters: a thematic role, which was always set to *agent* or *theme* for the examples we considered, and a level of granularity, which was assumed to be either atomic or provided by context. I assume that the thematic role parameter can also be set to τ , or runtime. Given that we take time to be a nonatomic domain, setting the thematic role parameter to τ should be incompatible with setting the granularity parameter to *Atom*. To put

it differently, there is no atomic-level distributivity operator for time. Only the Part operator can distribute over time.

Let us instantiate the Part operator in (41) with suitable dimension and granularity parameters. I will instantiate the dimension parameter as τ .

(70) **Definition: Event-based temporal Part operator**

$$\llbracket \text{Part}_{\tau, C} \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in {}^* \lambda e' \left(\begin{array}{c} V(e') \wedge \\ C(\tau(e')) \end{array} \right)$$

(Takes an event predicate V and returns a predicate that holds of any event e which consists entirely of events that are in V and whose runtimes satisfy the contextually salient ‘cover predicate’ C .)

The insertion of this temporal instantiation of Part can be seen as a repair strategy that shifts the meaning of a predicate in order to satisfy the atelicity presupposition of a *for*-adverbial. For example, *take a pill* by itself is punctual and telic, an achievement predicate in the terminology of Vendler (1957). Even lexical cumulativity does not change this fact, since no two such events can involve the same pill. But after the operator in (70) has been applied, the shifted predicate is now iterative. Depending on the value to which the variable C of this operator is resolved, this shifted predicate can be roughly paraphrased as *take a pill every day*, or *take a pill every hour*, etc., except that such overt quantifiers lack the context sensitivity of (41) and come with a consecutivity requirement I will discuss shortly. This kind of predicate can be empirically and formally shown to be atelic. Empirically, these predicates are atelic because they are compatible with *for*-adverbials, as shown by examples like (71), adapted from (56a).

(71) The patient took a pill every day for a month.

Formally, under an appropriate definition of atelicity formulated in terms of stratified reference, as in (22), it can be shown that any predicate that results from the application of the Part operator is atelic (see the Appendix). This means that when atelicity is formulated as in (22), an atelic predicate must have exactly the properties that the output of the temporal Part operator in (70) has, so long as the cover predicate is fine-grained enough.

We are now ready to explain the difference between examples (50a) and (58), repeated here:

(72) **Uttered out of the blue:**

??John found a flea for a month.

(73) **Context: discussing daily pill intake**

The patient took two pills for a month and then went back to one pill.

In example (73), I assume that the verb phrase *take two pills* has the following denotation prior to the application of the distributivity operator:

- (74) $\llbracket \text{take two pills} \rrbracket$
 $= \lambda e [* \text{take}(e) \wedge * \text{pill}(* \text{theme}(e)) \wedge |* \text{theme}(e)| = 2]$

This predicate applies to events in which a total of two pills are taken. It cannot combine directly with the *for*-adverbial because it is not atelic. The operator in (70) can be used as a repair strategy, provided that it is available. In (73), this is the case, because there is a salient level of granularity in the context and *C* can be resolved to it. That is, (73) is uttered in a context that makes salient a set of intervals that are no longer than a day—either the set $\lambda t. \text{days}(t) \leq 1$ or one of its subsets. This is as opposed to (72), which is uttered in a default context that does not contain anything that *C* can be resolved to.

This explains the contrast between (72) and (73). But how does (73) acquire its meaning? I have suggested that it involves the application of $\text{Part}_{\tau, \lambda t. \text{days}(t) \leq 1}$ to its verb phrase. The result of this operation is given in (75).

- (75) $\text{Part}_{\tau, \lambda t. \text{days}(t) \leq 1} (\lambda e [* \text{take}(e) \wedge * \text{pill}(* \text{theme}(e)) \wedge |* \text{theme}(e)| = 2])$
 $= \lambda e. e \in * \lambda e' \left(\begin{array}{l} * \text{take}(e') \wedge * \text{pill}(* \text{theme}(e')) \wedge \\ |* \text{theme}(e')| = 2 \wedge \\ \text{days}(\tau(e')) \leq 1 \end{array} \right)$
 (True of any plural event that consists of one or more events of taking two pills which each take place within a day.)

This predicate combines with *for a month* and then with *the patient*. Under plausible assumptions about the compositional process (see Section 2.10), and omitting the atelicity requirement for clarity, the result is the following:

- (76) $\exists e : \text{regular}(\tau(e)). * \text{agent}(e) = \text{the.patient} \wedge \text{months}(\tau(e)) = 1 \wedge$
 $e \in * \lambda e' \left(\begin{array}{l} * \text{take}(e') \wedge * \text{pill}(* \text{theme}(e')) \wedge \\ |* \text{theme}(e')| = 2 \wedge \\ \text{days}(\tau(e')) \leq 1 \end{array} \right)$
 (There is a regular plural event that consists of one or more events of taking two pills which each take place within a day. Its agent is the patient, and its runtime measures a month.)

This formula is verified by a regular plural event that unfolds over the course of a month. The subformula $\text{days}(\tau(e')) \leq 1$ makes sure that each of the taking-two-pills events e' takes place within a day. The pairs of pills covary with the days, even though there is no predicate such as *every day* explicitly mentioned in the sentence.

While I have not specified the pragmatic mechanism that supplies the value of the granularity parameter of *Part* in (75), it is instructive to try substituting various predicates for its actual value $\lambda t. \text{days}(t) \leq 1$. Some potential values are unavailable for various reasons. For example, the predicate *Atom* is not available in the temporal domain because, as mentioned earlier, natural-language semantics does not have

access to temporal atoms (see Section 2.4.4). Other potential values are unavailable because they denote times which are too long (in relation to the interval introduced by the *for*-adverbial) to satisfy the presuppositions of the *for*-adverbial (see Section 4.5.3). Finally, predicates that are not salient in the given context will be unavailable because the granularity parameter is taken to be anaphoric (Schwarzschild 1996).

To be sure, sentence (73) conveys more than the literal truth conditions expressed in (76). For example, while (76) is true in a scenario where the patient's daily pill intake is more than two, (73) typically conveys that the pill intake is exactly two. I assume that this is a scalar implicature. When generated in the scope of universal quantifiers and related operators, implicatures raise various theoretical issues; for an overview and discussion, see Schlenker (2016: sect. 22.2). Likewise, (73) implicates that there were no days at which the patient took no pills. This too is arguably an implicature since for some speakers it can be canceled, while inserting *every day* leads to a contradiction:

- (77) John used to take one pill at dinnertime. Then his prescription was doubled, and . . .
- a. . . he took two pills for a month, but he didn't stick to this regimen every day, so he didn't recover as quickly as he should have.
 - b. . . #he took two pills every day for a month, but he didn't stick to this regimen every day, so he didn't recover as quickly as he should have.

While judgments are variable, this contrast points to a possible difference between the Part operator and overt distributing expressions like *every day*. These expressions appear to have stronger truth conditions than Part in that they entail that relevant events occur on consecutive days, while Part only implicates this. Accordingly, the truth conditions in (76) do not require consecutivity, only regularity. I come back to the semantics of *every day* at the end of this section and in Section 9.5.

The strategy I have adopted here—Theory B plus a distributivity operator anaphoric on context—links the covariation ability of indefinites in the scope of *for*-adverbials to the presence of a salient temporal predicate. As we have seen, previous classes of theories lack this link. Theory A type accounts place no constraints on covariation, while theory B type accounts prevent it altogether.

The present system does not predict that all quantifiers take wide scope over *for*-adverbials. We have seen above that bare noun phrases do not take wide scope with respect to *for*-adverbials, because they denote algebraically closed predicates. This is illustrated in the following minimal pair, repeated here from (54a) and (72):

- (78) a. John found fleas (on his dog) for a month. OK
 b. ??John found a flea (on his dog) for a month. *odd out of the blue*

As discussed in Section 2.6.2, I assume that the denotation of a bare plural like *fleas* is the algebraic closure of its singular form, essentially *one or more fleas*. Due to lexical cumulativity, the bare plural in a predicate like *find fleas* stands in a cumulative-like

relation to each of the subintervals over which the *for*-adverbial quantifies. Sentence (78a) does not entail that any one flea has been found several times, only that there is a plural month-long interval over the course of which one or more fleas were found. This is entailed by the following representation of the denotation of (78a), where *fleas* is interpreted in situ as a predicate that applies to the theme of the verb:

- (79) $\llbracket \text{John found fleas for a month} \rrbracket$
 $= \exists e : \text{regular}(\tau(e)) \wedge \text{SR}_{\tau, \gamma(\tau, e)}(\lambda e' [*find(e') \wedge *flea(*theme(e'))])(e).$
 $[*agent(e) = j \wedge *find(e) \wedge *flea(*theme(e)) \wedge \text{months}(\tau(e)) = 1]$
 (Presupposes that there is a regular event that is stratified by the property of finding a set of fleas. Asserts that this event is a finding event whose theme is a set of fleas, whose agent is John, and whose runtime measures one month.)

The definedness condition of this sentence is fulfilled since *find fleas* has stratified reference with respect to time (see Chapters 4 through 6). Given lexical cumulativity, any event in virtue of which (78a) is true can consist of several individual findings. No distributivity operator is needed. This observation is parallel to the argument about meaning postulates in Lasersohn (1989). I discussed this argument in Section 8.3 in connection with sentences like (15) and (20), repeated here as (80a) and (80b):

- (80) a. Rodgers, Hammerstein, and Hart wrote musicals. *nonatomic*
 b. Rodgers, Hammerstein, and Hart wrote a musical. **nonatomic*

The system presented here explains the contrast in (80) in the same way as the contrast in (78). In both cases, lexical cumulativity causes the sentence with the bare plural object to exhibit nonatomic lexical distributivity. And in both cases, the sentence with the singular indefinite object does not exhibit nonatomic phrasal distributivity because the lack of supporting context means that the nonatomic distributivity operator *Part* is not available.

The idea that covariation of singular indefinites is due to the insertion of a covert operator finds additional support in the observation that covarying singular indefinites take extra time to process compared with bare plurals, as discussed in Deo & Piñango (2011). Reading time increases at the *for*-adverbial in (81a) compared with (81b) in self-paced reading tests conducted by Todorova, Straub, Badecker & Frank (2000). The indefinite in (81a) is able to covary, that is, it involves reference to more than one large check.

- (81) a. Even though Howard sent a large check to his daughter for many years, she refused to accept his money.
 b. Even though Howard sent large checks to his daughter for many years, she refused to accept his money.

On the view presented here, example (81a) involves the covert presence of a *Part* operator but example (81b) does not. This kind of contrast can be explained if we

assume that the retrieval of an antecedent for the anaphoric variable *C* in the Part operator leads to a higher processing load. I take examples like (81a) to show that anaphoricity is built into the distributivity operator rather than into the *for*-adverbial, pace Deo & Piñango (2011).

I have not yet explained why overt interveners like *every day* can cause singular indefinites in the scope of *for*-adverbials to involve reference to multiple entities, as in these examples, repeated here from (56):

- (82) a. John found a flea on his dog every day for a month.
 b. Jim hit a golf ball into the lake every five minutes for an hour.

The answer is simple: It is the overt universal quantifier in these sentences that causes indefinites in its scope to covary with it, and not the *for*-adverbial. (The indefinite can in turn escape the scope of the universal quantifier by taking scope above the sentence; the resulting reading is plausible in (82b) but not in (82a).)

This universal quantifier plays a similar role to the Part operator in (70). For this reason, and for other reasons discussed in Section 9.5, I propose to analyze *every day* analogously to that operator, with the following modifications (and the analysis of *every five minutes* is similar). First, *every day* hardwires the granularity parameter to the value $\lambda t[\text{days}(t) \leq 1]$ instead of retrieving it anaphorically from context. Second, as we have seen, *every day* entails that no relevant day is left out, while the Part operator merely implicates this.

The output of *every day* must satisfy the atelicity requirement imposed by the *for*-adverbial. This would not be the case, for example, if we let *every day* universally quantify over a fixed set of days. The underlined formula in (83) ensures that every day in some consecutive span is the runtime of a relevant event, and it satisfies the atelicity requirement by existentially quantifying over that span rather than fixing it.

- (83) $\llbracket \text{every day} \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e [e \in * \lambda e' \left(\frac{V(e') \wedge \text{days}(\tau(e')) \leq 1}{\exists t. \text{continuous}(t) \wedge \forall t'. [t' \leq t \wedge \text{days}(t') = 1 \rightarrow \exists e'. e' \leq e \wedge V(e') \wedge \tau(e') \leq t']} \right) \wedge$
 (Takes an event predicate *V* and returns a predicate that holds of any event *e* which consists entirely of events that are in *V* and whose runtimes are at most one day long, in such a way that no day in a consecutive span of days is left out.)

While this entry induces covariation in indefinites as desired, it is only a sketch and raises avenues for further research. For example, one could try to derive the requirement that days be consecutive from general properties of domain restriction (see e.g. Stanley & Szabó 2000; Schwarz 2009: ch. 3). Those properties might perhaps also capture the fact that *every day* quantifies over a set that is maximal in some salient way but that can be bounded by temporal adverbials and tense. For example, in (82a), *every day* quantifies over some set of past days that are all contained within the same

month. A proper account of *every day* must be integrated into an account of temporal dependencies (Pratt & Francez 2001; von Stechow 2002b; Champollion 2011b). Furthermore, making (83) compositional would need more work, whether we build on classical accounts of *every* or on the event-based treatment in Section 9.5. Either way, one would need to determine the source of the consecutivity requirement of *every day*, which has no obvious counterpart in ordinary cases of universal quantification such as *every dog*. The behavior of pluractional adverbials like *day after day*, *dog after dog*, etc. might be relevant here (Beck & von Stechow 2007).

To sum up this section, there are two kinds of theories of the scopal effects of *for*-adverbials on indefinites: those that predict indefinites can always covary, and those that predict that they never do. The first kind of theory overgenerates and the second kind undergenerates. I have solved this puzzle by transferring the cover-based approach to distributivity in Schwarzschild (1996) into the temporal domain. Adopting Neo-Davidsonian event semantics with its parallel treatment of arguments and adjuncts made this transfer easy. On the resulting account, it is not the *for*-adverbials that induce covariation but contextually dependent distributivity operators. These operators are similar in this respect to overt quantifiers like *every day*, which induce covariation when they intervene between the *for*-adverbial and the indefinite.

8.7 Summary

This chapter has presented a theory of covert phrasal distributivity in natural language. Covert phrasal distributivity has been at the center of a long debate as to whether it always involves distribution over atoms (Laserson 1989; Winter 2001) or whether it can also involve distribution over pragmatically salient nonatomic entities (Gillon 1990). The framework deployed here combines ideas from algebraic semantics (Link 1983) and Neo-Davidsonian event semantics (Parsons 1990), which exposes thematic roles to the compositional process. I have suggested that distributivity operators are parametric on thematic roles and, in certain cases, on salient predicates such as covers. I have suggested two divisions of labor, one between meaning postulates and distributivity operators and the other one between semantics and pragmatics, that account for the limited availability of nonatomic distributivity. This provides an advantage over theories in which nonatomic distributivity is freely available, those in which it is not available at all, and those where its availability does not depend on the difference between lexical and phrasal distributivity. I have extended this account to the temporal domain, where it predicts the limited availability of indefinites to covary in the scope of *for*-adverbials. This provides an advantage over theories that predict such indefinites to always covary and those that predict them to never covary. I have reformulated the atomic distributivity operator *D* of Link (1987b) and the cover-based nonatomic distributivity operator Part of Schwarzschild (1996) in algebraic

event semantics in a way that makes the sum event available for further modification by arguments and adjuncts, which was not the case in previous implementations.

Like stratified reference, these distributivity operators have two parameters. The first parameter specifies the dimension or thematic role that is targeted, which may also be runtime. The second parameter specifies the granularity or size of the distributed entities. The connection between the distributivity operators and stratified reference allows us to connect the present theory of distributivity more broadly to aspect and measurement (see Chapter 4).

I have empirically distinguished lexical and phrasal distributivity. Following earlier work, I have suggested that lexical distributivity should be modeled by the lexical cumulativity assumption (Lasersohn 1989; Kratzer 2007). In keeping with the observation that nonatomic phrasal distributivity exists but requires supporting context to surface, I have suggested that this fact motivates the possibility of a nonatomic granularity parameter setting for the distributivity operator. I have assumed that this happens so rarely because the operator is anaphoric on its context with respect to this parameter (Schwarzschild 1996). Examples like (21), repeated as (84a), have a nonatomic reading because there is a salient nonatomic level of granularity, while examples like (23), repeated as (84b), do not have such a reading. I have extended this parallel to the temporal domain, where I have argued that a salient level of granularity provides a way for the indefinite in (73), repeated as (85a), to involve reference to different sets of two pills, while such a reading is not present in (85b).

- (84) a. The shoes cost fifty dollars (*i.e. per pair*).
b. The suitcases weigh fifty pounds.
- (85) a. The patient took two pills for a month (*i.e. per day*).
b. ??John found a flea for a month (*i.e. per day*).

With this theory of covert distributivity in hand, we are ready to turn to overt distributivity. Across languages, overt phrasal distributivity is often expressed via adverbials and adnominals, such as English *each* and German *jeweils* (Zimmermann 2002b). Such items differ with respect to whether they are restricted to distribution over individuals mentioned in the same sentence (like D), or whether they can also distribute over pragmatically salient occasions that need not have been explicitly mentioned (like Part). This parallel to the D and Part operators is central to Chapter 9.

Overt distributivity

9.1 Introduction

This chapter presents a compositional theory of distance distributivity that relates adnominal and adverbial distributive items, atomic and cover-based distributivity, and distributive determiners to each other. It has been previously published as Champollion (2016c) and is reprinted here with slight modifications. While Chapter 8 focused on covert distributivity, this chapter is about overt distributivity. Overt and covert distributivity are illustrated in (1).

- (1) a. The girls each wore a black dress.
b. The girls wore a black dress.

In sentence (1a), the adverbial distributive item *each* distributes the predicate *wear a black dress* over the individual girls and leads to the entailment that each of the girls in question wears a black dress. Sentence (1b) is interpreted in the same way, even though there is no *each*. As discussed in Section 8.2, the ability of verb phrases to distribute in the absence of an overt distributive item has been attributed to the D operator.

The purpose of this chapter, along with chapter 8, is to bring together several strands of research on phenomena related to the semantics and pragmatics of distributivity in natural language. One of these strands deals with overt distributivity, which is crosslinguistically often expressed via adverbials and adnominals, such as English *each* and German *jeweils* /'je:vails/. Such elements differ with respect to whether they are restricted to distribution over individuals mentioned in the same sentence, or whether they can also distribute over pragmatically salient occasions that need not have been explicitly mentioned (Moltmann 1997). This strand is motivated by the properties of adverbial *each* and its adnominal and determiner counterparts both in English and in other languages (Zimmermann 2002b). As we will see, the meanings of these elements vary in ways that sometimes require them to distribute over atoms (individuals), such as in the case of *each*, and in other cases also allow them to distribute over salient nonatomic entities, such as occasions in the case of *jeweils*, as illustrated here:

- (2) Hans hat jewels zwei Affen gesehen.
Hans has DIST two monkeys seen.
'Hans has seen two monkeys on each occasion.' (German)

Another strand concerns the properties of silent distributivity operators such as the one arguably present in (1b). As discussed in Section 8.1, covert phrasal distributivity has been at the center of a long debate as to whether it always involves distribution over atoms—singular individuals—or whether it can also involve distribution over nonatomic entities (Lasersohn 1989, Gillon 1990). I reserve the term *D operator* for distributivity operators that always distribute over atoms. As for the nonatomic version of the operator, whose meaning may be paraphrased as *each salient part of*, I will refer to it as the *Part operator*, following Schwarzschild (1996) (see Section 8.3).

As this sketch already suggests, overt and covert distributivity share many similarities. In both cases, some elements can only distribute to atoms (*each*, *D*) while others can distribute to salient nonatomic entities (*jewels*, *Part*). And as we will see, in both cases the former elements can only distribute over pluralities that have been explicitly mentioned, while the latter elements can also distribute over salient domains that have not been explicitly mentioned, such as temporal occasions. These similarities give rise to analogous questions in the overt and in the covert case. Can a given distributive item (be it a covert operator or a word) only distribute down to singular entities or also to plural entities? Do these entities need to be of a certain size or “granularity,” and can this size vary from element to element? Must these entities have been overtly mentioned in the sentence and thereby contributed by syntactic means, or can they also be supplied by the context via pragmatic means?

A unified semantic analysis of distributivity should make it apparent which aspects of the meanings of various distributivity operators are always the same, and along which dimensions these meanings can differ. The theory should capture the semantic variation across distributivity-related elements. The resulting system should be fully formalized and explicit.

This chapter, together with Chapter 8, contributes towards these goals. By combining ideas from algebraic semantics and event semantics, we can provide a framework in which the split in overt distance-distributive items can be related to the debate in the literature on covert distributivity. In this framework, the various uses of *each* and similar items in other languages are all related to the distributivity operator, either in its semantic, atomic form as defined by Link (1987b) or in its pragmatic, salience-related form as defined by Schwarzschild (1996). As we will see, these various uses of *each* and these silent operators share some part of their meanings with each other and with their counterparts across languages. This fact is captured by deriving them from related distributivity operators which differ only in possible settings of the two strata-theoretic parameters and the ranges of values they allow for them. One parameter indicates the dimension in which distributivity takes place. This can be a thematic role in some semantic instances of distributivity, or a spatial or temporal dimension in other instances. The other parameter indicates the size of the entities over which distributivity takes place, such as atoms or salient amounts of space or time. The two parameters interact with each other against the background of assumptions about the

metaphysics of natural language. For example, time is assumed to be either nonatomic or in any case to not make its atoms available to the semantics of natural language. As a result, when the first parameter is set to time, the second cannot be set to anything involving atoms, because time does not provide any atoms to distribute over. The analysis in this chapter is compositional and avoids unusual semantic concepts such as index-driven and crosswise λ -abstraction (Zimmermann 2002b) or distributive polarity items (Oh 2001, 2006).

As we saw in Section 8.4, the Neo-Davidsonian event-semantic setting gives us the ability to think of the D and Part operators as being coindexable with different thematic roles. This allows us to capture through a simple change in coindexation the kinds of configurations that have otherwise been taken to require type-shifting-based reformulations of these operators (Lasnik 1998b):

- (3) a. The first-year students [D [took an exam]]. *Target: agent*
 b. John [D [gave a pumpkin pie]] to two girls. *Target: goal*

The reformulation of the distributivity operators in Section 8.4 provides the groundwork on which I build in this chapter in order to formally relate this ambiguity to the one observable in examples like these:

- (4) The boys told the girls two stories each. *Target: agent*
 (two stories per boy)
 (5) The boys told the girls two stories each. *Target: goal*
 (two stories per girl)

To capture this and other parallels between covert and overt distributivity, I will propose that distance-distributive items across languages are in essence overt versions of the D and Part operators.

Strata theory provides us with a way to formulate commonalities and differences across instances of distributivity in natural language. Individual elements can be analyzed as being hardwired for certain parameter values, so that, for example, the difference between Link's and Schwarzschild's operators, as well as that between *each* and *jewels*, can be described in terms of whether the value of the granularity parameter is prespecified to *Atom* or can be filled in by context. In this way, overt and covert instances of distributivity fit together and into distributivity theory more generally.

To develop this picture, Section 9.2 starts by describing relevant facts and generalizations about overt instances of distributivity across languages, drawing largely on the crosslinguistic discussion in Zimmermann (2002b). Overt and covert distributivity are brought together in Section 9.3, which develops a compositional account of overt distance distributivity. Section 9.4 deals with more complicated syntactic configurations, some of which previously lacked a compositional analysis. Section 9.5 extends the analysis to the determiners *each* and *every*, and shows that it accounts for

their ability to take part in cumulative readings and to interact with nondistributive event modifiers. The way in which meanings of overt distributive items vary across languages is explained in Section 9.6. Section 9.7 compares the present analysis with Zimmermann (2002b), and Section 9.8 concludes.

9.2 Overt distributivity across languages

Distributive items have different syntactic uses and different meanings across languages. In English, the distributive quantifier *each* can be used in at least three ways. This is shown in the examples in (6), repeated from Section 3.4:

- (6) a. **Adnominal:** Two men have carried three suitcases *each*.
- b. **Adverbial:** Two men have *each* carried three suitcases.
- c. **Determiner:** *Each* man has carried three suitcases.

There are many terms for these three uses. Adnominal *each* has also been called shifted (Postal 1974), an anti-quantifier (Choe 1987), binominal (Safir & Stowell 1988), or ditransitive (Roberts 1987). Adverbial *each* has also been called floated (Choe 1987). Determiner *each* is also called prenominal (Safir & Stowell 1988). For the purpose of this chapter, I set aside the use of *each* in other constructions, notably the reciprocal *each other* and the partitive *each of the men*. LaTerza (2014a, 2014b) connects all these constructions in a similar framework to the one I adopt here. I will refer to the noun phrase *two men* in (6a) and (6b) as the antecedent (or target) of *each*, and to the noun phrase *three suitcases* in (6a) as the host of adnominal *each*.

I will refer to adnominal and adverbial *each* and to similar elements across languages as *distance-distributive* items. That term is taken from Zimmermann (2002b). There is a slight difference in terminology: Zimmermann reserves the term *distance distributivity* for adnominal elements, while I use it both for adnominal and for adverbial elements. This seems appropriate because adverbial *each* can be separated from its antecedent, for example by an auxiliary as shown in (6b).

Adnominal *each* can be shown by movement tests to form a constituent with its host noun phrase (Burzio 1986, Safir & Stowell 1988). Distance-distributive items like it are sometimes seen as a challenge for compositional semantics, because their interpretations are similar to those of distributive determiners even though their surface syntactic structure appears to be fundamentally different (Oh 2001, 2006). For example, adnominal *each* in the object of sentence (6a) is contained in the constituent over which it seems at first sight to distribute, namely the verb phrase *carried three suitcases each*. This is of course similar to the challenge represented by quantifiers in object position (*carried every suitcase*), and the standard solutions to that challenge are available in both cases. For example, one can lift the type of the quantifier or the verb in order to give the quantifier scope over the verb phrase (Hendriks 1993, Barker 2002). I will adopt the same general strategy in the formal analysis which follows. In fact, we will see that the scope of adnominal *each* is even

more restricted than that of object quantifiers, because it does not include the verbal projection.

As we saw in (4) and (5), adnominal *each* can target different antecedents. This dependency is generally regarded as a case of ambiguity rather than underspecification. The ambiguity view finds support in analogous constructions involving dependent numerals in American Sign Language, where anaphoric dependencies are realized overtly (Kuhn 2015). Further support comes from the fact that in some languages, adnominal distance-distributive items must agree with their antecedents. There are syntactic constraints on the distribution of adnominal *each* with respect to its antecedent, such as c-command requirements and clausemate conditions. Accordingly, the dependency has been variously argued to be similar to that of reflexive pronouns with respect to their antecedents (Burzio 1986, Safir & Stowell 1988) or to that of traces of noun phrases that undergo raising with respect to these noun phrases (Sportiche 1988). Similarly, adverbial *each* has been variously claimed to be related to its antecedent by movement, in the sense that it modifies the trace of its antecedent, or to be base-generated, in which case its relation to its antecedent can be taken to be anaphoric. For an overview of these conflicting claims and their implications, see Bobaljik (2001).

I will not add to the discussion on these syntactic constraints. Since the nature of the dependency between adnominal and adverbial *each* and their antecedents is not the focus of this chapter, I will not take a strong position on it. In the formal theory to be developed in this chapter, I will represent it by coindexing distance-distributive items with thematic roles, in the same way as I have coindexed covert distributivity operators with thematic roles in Section 8.4. More specifically, as discussed in Section 2.10, I will assume that thematic roles are introduced into the compositional derivation by θ -role heads, and that distance-distributive items in certain languages including English must be coindexed with a θ -role head in their clause. Accordingly, I will refer to this coindexation as θ -indexing.

I assume that θ -indexing is similar to other dependencies that are commonly formalized by coindexation, for example those between reflexive pronouns and their antecedents, in that it is subject to binding-theoretic constraints (Chomsky 1981, Buring 2005). In assuming that θ -indexing is binding-theoretic in nature, I follow previous semantic analyses that interpret adnominal *each* without any movement, such as Zimmermann (2002b). Should the movement-based view turn out to be the correct one instead, the syntax-semantics interface may need to be modified accordingly, for example by incorporating elements from the theory of Cable (2014).

Turning now to other languages, adnominal and adverbial *each* can be translated in German by the word *jeweils* (Moltmann 1997, Zimmermann 2002b). Determiner *each*, however, must be translated by another word, *jeder*. Here and throughout this chapter, I gloss distance-distributive items as DIST rather than EACH or EVERY since in some cases they have a wider range of readings than *each*. Sometimes they mean *each*, sometimes their meaning is closer to *each time* or *on each occasion*.

Example (7a) is adnominal, example (7b) is adverbial, and example (7c) contains a determiner. Though adverbial and adnominal *jeweils* take the same surface position in (7a) and (7b), they can be teased apart syntactically, as discussed in Zimmermann (2002b).

- (7) a. Die Jungen haben [*jeweils* [drei Koffer]] getragen.
 The boys have DIST three suitcases carried
 'Each of the boys has carried three suitcases.' / *In certain contexts*:
 'The boys have carried three suitcases each time.'
- b. Die Jungen haben [*jeweils* [drei Koffer getragen]].
 The boys have DIST three suitcases carried
 (As above.)
- c. Jeder/**Jeweils* Junge hat drei Koffer getragen.
 Each.sg.m/DIST boy has three suitcases carried
 'Each boy carried three suitcases.' (German)

As we will see, *each* and *jeweils* generalize to two classes of distance-distributive items across languages. *Each*-type distance-distributive items can also be used as determiners. *Jeweils*-type distance-distributive items cannot double as determiners, as shown in (7) for *jeweils* itself. Some languages have distance-distributive items which can also function as distributive determiners, as in English, and others are like German in that they have no such elements (Zimmermann 2002b). Partly building on earlier typological work (e.g. Gil 1982), Zimmermann observes that distance-distributive items which can also be used as determiners (e.g. *each*) always distribute over individuals, as determiners do. In contrast, those distance-distributive items which are formally distinct from determiners can typically also distribute over salient occasions, that is, over chunks of time or space.

The range of crosslinguistic variation can be illustrated by comparing English *each* with German *jeweils*, a distance-distributive element which cannot double as a distributive determiner. *Jeweils* can distribute over individuals like English *each*, but also over spatial or temporal occasions, as long as context provides a salient set of such occasions (Gil 1993, 1995, Moltmann 1997, Zimmermann 2002b). I call this the *occasion reading*. It corresponds to what is also called the *temporal key reading* and the *spatial key reading* (Balusu 2005, Balusu & Jayaseelan 2013). I will treat the temporal and spatial cases as two separate readings (and I will focus on the temporal case). I leave the question open whether they should be treated as special cases of one reading. Another, less theory-neutral term for the occasion reading is *event-distributive reading* (Oh 2001, 2006). Zimmermann (2002b) uses the term *adverbial reading* for it. This term is potentially misleading, because it suggests that only the adverbial use of *jeweils* can give rise to this reading. But adnominal *jeweils* can give rise to it as well (Zimmermann 2002b: ch. 5). For example, in (8), *jeweils* is part

of the subject noun phrase⁴³ and is therefore adnominal. However, as shown by the paraphrase, this instance of *jeweils* distributes over occasions, not over individuals.

- (8) *Jeweils zwei Jungen standen Wache.*
 DIST two boys stood watch
 ‘Each time, two boys kept watch.’ (German)

The examples in (9) illustrate the occasion reading. Sentence (9) is ambiguous between a reading that distributes over individuals—the ones of which their plural subject consists, (9a)—and one that distributes over occasions (9b).

- (9) *Die Jungen haben jeweils zwei Affen gesehen.*
 The boys have DIST two monkeys seen
 a. ‘Each of the boys has seen two monkeys.’
 b. ‘The boys have seen two monkeys each time.’ (German)

While the former reading is always available, the latter requires a supporting context. That is, when (9) is uttered out of the blue, it only has the reading (9a). The occasion reading (9b), by contrast, is only available in contexts where there is a previously mentioned or otherwise salient set of occasions, such as contexts in which the boys have been to the zoo on several previous occasions.

Unlike *each*, *jeweils* can also occur with a singular subject, as in (10), repeated here from (2), which only has an occasion reading.

- (10) *Hans hat jeweils zwei Affen gesehen.*
 Hans has DIST two monkeys seen
 ‘Hans has seen two monkeys on each occasion.’ (German)

This sentence is odd out of the blue, and it requires supporting context in the same way as reading (9b) does. Its other potential reading would involve vacuous distribution over only one individual, Hans. This is presumably blocked through the Gricean maxim of manner “Be brief” or a non-vacuity presupposition or implicature or whatever else prevents vacuous distributivity (Roberts 1987: 219). For more on this point, and for a fuller discussion of the kinds of noun phrases that can license adnominal *each* in English, see Champollion (2015a).

In the presence of contextual cues, *jeweils* is also able to distribute over nonatomic entities such as contextually salient groupings of atomic individuals:

- (11) *Beim Zoobesuch wurden die Jungen in Fünfergruppen aufgeteilt.*
 At.the zoo.visit were the boys in quintuplet.groups divided
Die Jungen sahen jeweils zwei Affen.
 The boys saw DIST two monkeys

⁴³ We know this because German as a V2 language allows only one constituent before the tensed verb *standen*.

'During the zoo visit, the boys were divided up into groups of five. Each of the boys (or: of the groups) saw two monkeys.' (German)

In this example, *jeweils* can be understood in two ways. Depending on how this ambiguity is resolved, the truth conditions require two monkey sightings per boy or per group of five boys. As discussed in Section 8.3, I follow Schwarzschild (1996) in treating such pluralities as salient nonatomic entities, rather than as "group atoms" in the sense of Landman (1989). This allows me to treat *jeweils* on par with the Part operator in Schwarzschild (1996), which also requires salience when it distributes over nonatomic entities. That operator is anaphoric on a contextually salient cover over the plurality in question, that is, a set of salient (and potentially plural) entities whose sum is that plurality.

The previous example suggests that a grouping of contextually salient pluralities (or occasions) is sufficient for nonatomic interpretations of *jeweils*. In fact, it is a prerequisite. Out of the blue, *jeweils* is not able to distribute over nonatomic entities. This can be seen in the sentence in (12), which is based on an example in Gillon (1987):

- (12) Rodgers, Hammerstein und Hart haben jeweils ein Musical
Rodgers, Hammerstein and Hart have DIST a musical
geschrieben.
written.

'Rodgers, Hammerstein and Hart each wrote a musical.' (German)

Out of the blue, (12) entails that each of the three composers wrote a musical, and cannot be interpreted as true in a scenario where Rodgers and Hammerstein wrote a musical together, Rodgers and Hart wrote another musical together, and no other musicals were written (see Section 8.3 for discussion of this scenario). In other words, the mere existence of a nonatomic cover is not sufficient to make a *jeweils* sentence true. If (12) is uttered out of the blue, such a cover is not salient, and the putative reading on which the sentence would be true does not arise.

Given an appropriate context, *jeweils* can also distribute within the mass domain. This can be seen in (13), a variant of an example discussed in connection with the Part operator in Section 8.3:

- (13) Auf der Party wurden Milch und Limonade angeboten. Milch kostete
At the party were milk and lemonade offered. Milk cost
jeweils fünfzig Cent und Limonade kostete jeweils einen Euro.
DIST fifty cent and lemonade cost DIST one euro.
'At the party, they sold milk and lemonade. Milk cost fifty cents, and lemonade
cost one euro.' (German)

The predicates in question are distributed to a contextually salient level, namely the units in which the beverages are sold. (The last three examples were suggested by a reviewer for the journal *Semantics and Pragmatics*.)

While *jeweils* allows distribution both over individuals and over salient nonatomic entities, this is not the case for all distance-distributive items (Zimmermann 2002b). Across languages, many adnominal distance-distributive items can only distribute over individuals. For example, English adnominal *each* lacks the occasion reading:

- (14) The boys saw two monkeys each.
 a. *Available*: 'Each of the boys saw two monkeys.'
 b. *Unavailable*: 'The boys saw two monkeys on each occasion.'

When adnominal *each* is used in a sentence whose subject is singular, distribution over individuals is not possible, presumably for pragmatic reasons as mentioned:

- (15) *John saw two monkeys each.

Unlike (10), this sentence lacks an occasion reading, even with supporting context. To make the occasion reading surface, one cannot just provide temporal antecedents, as in (16a). Instead, one must add an overt noun like *time* as a complement of *each*, as in (16b).

- (16) a. *Yesterday and today John saw two monkeys each.
 b. John saw two monkeys each time.

We have seen that English *each* also differs from German *jeweils* in that only the former can also be used as a determiner. Coming back to what I mentioned at the beginning of this section, Zimmermann postulates the following crosslinguistic generalization (Zimmermann 2002b):

(17) **Zimmermann's generalization**

All *each*-type distance-distributive items (i.e. those that can also be used as determiners) can only distribute over individuals. This contrasts with *jeweils*-type distance-distributive items, many of which can also distribute over salient spatial or temporal occasions.

This generalization is based in part on the following examples, which show that Albanian, Dutch, French, Icelandic, Italian, Japanese, Norwegian, Portuguese, Russian, and possibly Latin all have distance-distributive items that behave like English *each* in two ways: They can also be used as distributive determiners, and they lack the occasion reading, except in some cases when an extra noun with the meaning *time* (in the sense of *occasion*) is added. Many of the following examples are from Zimmermann (2002b).

- (18) Fëmijët blenë secili dy salsiçe.
 Children.def.nom buy.aor.3pl DIST.nom two sausages
 'The children bought two sausages each.' (Albanian)⁴⁴

⁴⁴ Examples (18)–(20): Bujar Rushiti, p.c. to the author.

- (19) *Secili fëmijë ka blerë nga dy salsiçe.*
 DIST.nom child.sg.nom has bought DIST two sausages
 'Each child bought two sausages.' (Albanian)
- (20) **Beni ka blerë secili dy salsiçe.*
 Ben.nom has bought DIST.nom two sausages
 Intended: 'Ben bought two sausages each time.' (Albanian)
- (21) *De jongens hebben elk twee boeken gelezen.*
 the boys have DIST two books read
 'The boys have read two books each.' (Dutch)⁴⁵
- (22) *Elke jongen heeft twee boeken gelezen.*
 DIST boy has two books read
 'Each boy has read two books.' (Dutch)⁴⁶
- (23) *Hans heeft elke *(keer) twee boeken gelezen.*
 Hans has DIST time two books read
 'Hans has read two books each time.' (Dutch)⁴⁷
- (24) *Les professeurs ont lu deux livres chacun/chaque.*
 the professors have read two books DIST
 'The professors have read two books each.' (French)⁴⁸
- (25) *Chaque professeur a lu deux livres.*
 DIST professor has read two books
 'Each professor has read two books.' (French)⁴⁹
- (26) *Pierre a lu deux livres { chaque *(fois) / *chacun(e) (fois). }*
 Pierre has read two books { DIST time / DIST time }
 'Pierre read two books each time.' (French)⁵⁰
- (27) *Strákarnir keyptu tvær pylsur hver.*
 the.boys bought two sausages DIST
 'The boys bought two sausages each.' (Icelandic)⁵¹

⁴⁵ Zimmermann (2002b: 40).⁴⁶ Zimmermann (2002b: 44); corrections supplied by Floris Roelofsen, p.c. to the author.⁴⁷ Floris Roelofsen, p.c. to the author.⁴⁸ Tellier & Valois (1993: 574, ex. 1a) quoted in Zimmermann (2002b: 41). *Chaque* is colloquial as an adnominal. While French adnominal *chacun* and determiner *chaque* are not exactly identical, they are historically related and can still be considered formally identical (Grevisse 1980, Junker 1995, Zimmermann 2002b: 44, n. 30).⁴⁹ Zimmermann (2002b: 44).⁵⁰ Author's judgment, adapted from Zimmermann (2002b: 47).⁵¹ Examples (27) through (29): Meike Baumann, Daði Hafþór Helgason, Hildur Hrólfssdóttir, Sverrir Kristinsson, Gunnar Ingi Valdimarsson, p.c. to the author.

- (28) *Hver strákur keypti tvær pylsur.*
 DIST boy bought two sausages
 'Each boy bought two sausages.' (Icelandic)
- (29) *Pétur keypti tvær pylsur hvert *(sinn).*
 Pétur bought two sausages DIST time
 'Pétur bought two sausages each time.' (Icelandic)
- (30) *I ragazzi comprarono un libro ciascuno.*
 the boys bought a book DIST.sg.m
 'The boys bought one book each.' (Italian)⁵²
- (31) *Ciascun ragazzo ha comprato due salsicce.*
 DIST.sg.m boy has bought two sausage.pl.f
 'Each boy has bought two sausages.' (Italian)⁵³
- (32) **Peter ha comprato due salsicce ciascun/-o/-e.*
 Peter has bought two sausage.pl.f DIST.sg.m/DIST.sg.m/DIST.pl.f
 Intended: 'Peter has bought two sausages.' (Italian)⁵⁴
- (33) *Otoko-tati-ga sorezore huta-ri-no zyosei-o aisi-teiru koto.*
 men-pl-nom DIST two-cl-gen women-acc love-asp fact
 'The fact that the men love two women each.' (Japanese)⁵⁵
- (34) *Sorezore-no gakusei-ga iti-dai-no piano-o motiage-ta.*
 DIST-gen student-nom one-cl-gen piano-acc lift-past
 'Each student lifted one piano.' (Japanese)⁵⁶
- (35) *Taroo-wa sorezore-?(de) iti-dai-no piano-o motiage-ta.*
 Taroo-top DIST(-loc) one-cl-gen piano-acc lift-past
 'Taroo lifted one piano on each occasion.' (Japanese)⁵⁷
- (36) *Guttene har kjøpt to pølser hver.*
 boys-the have bought two sausages DIST
 'The boys bought two sausages each.' (Norwegian)⁵⁸

⁵² Burzio (1986: 198, ex. 50b) quoted in Zimmermann (2002b: 41).

⁵³ Zimmermann (2002b: 44).

⁵⁴ Ivano Ciardelli, p.c. to the author.

⁵⁵ Sakaguchi (1998: 115, ex. 1), quoted in Zimmermann (2002b: 41).

⁵⁶ Sakaguchi (1998: ex. 7).

⁵⁷ Chigusa Kurumada, p.c. to the author. Kurumada comments that the sentence without *de* feels like an elliptical version of the sentence with *de*.

⁵⁸ Øystein Vangsnes, p.c. to Zimmermann (2002b: 40).

- (37) *Hver* gutt har kjøpt to pølser.
DIST boy has bought two sausages
'Each boy has bought two sausages.'
(Norwegian)⁵⁹
- (38) Jon har kjøpt to pølser *hver* *(gang).
Jon has bought two sausages DIST time
'Jon has bought two sausages each time.'
(Norwegian)⁶⁰
- (39) Os meninos compraram duas linguiças *cada* (um).
The boys bought two sausages DIST (one)
'The boys bought two sausages each.'
(Portuguese)⁶¹
- (40) *Cada* menino comprou duas linguiças.
DIST boy bought two sausages
'Each boy bought two sausages.'
(Portuguese)⁶²
- (41) João comprou duas linguiças *cada* *(vez).
João bought two sausages DIST time
'João bought two sausages each time.'
(Portuguese)⁶³
- (42) Mal'chiki kupili po dve sosiski *kazhdyj*.
boys.nom buy.pfv.past.pl DIST two sausage DIST
'The boys bought two sausages each.'
(Russian)⁶⁴
- (43) *Kazhdyj* mal'chik kupil dve sosiski.
DIST boy buy.pfv.past.m.sg two sausage
'Each boy bought two sausages.'
(Russian)⁶⁵
- (44) *Kazhdyj* *(raz) Petja pokupal dve sosiski.
DIST (time) Petja buy.ipfv.past.m.sg two sausage
'Each time, Petya bought two sausages.'
(Russian)⁶⁶

To this list we may tentatively add Latin *quisque* 'each, any,' which is described as lacking the occasion reading (Bortolussi 2013: 13, n. 11) and which can function both as a distance-distributive item and as a determiner:

⁵⁹ Zimmermann (2002b: 44).

⁶⁰ Kjell Johan Sæbø, p.c. to the author.

⁶¹ Examples (39) through (41): Luciana Meinking Guimarães and Leonor Remédio, p.c. to the author. This example is given in Brazilian Portuguese. In European Portuguese, *boy(s)* would be translated more commonly as *rapaze(s)*.

⁶² Brazilian Portuguese. See n. 62.

⁶³ Brazilian Portuguese. In European Portuguese, this would be translated as *O João comprou duas linguiças de cada* *(vez).

⁶⁴ Olga Borik, p.c. to Zimmermann (2002b: 41). On *po*, see also examples (63) through (65).

⁶⁵ Olga Borik, p.c. to Zimmermann (2002b: 44).

⁶⁶ Masha Esipova and Sonia Kasyanenko, p.c. to the author.

- (45) duo *quisque* Alpina coruscant / gaesa
 two.n DIST.nom.sg.m Alpine.nom.pl.n sparkle.3pl.sg / javelin.nom.pl.n
 manu
 hand.abl.sg.f
 ‘in each man’s hand two Alpine javelins gleamed’ (Latin)⁶⁷

- (46) quod *quisque* imperator habeat pecuniae
 what DIST.nom.sg.m general.nom.sg.m have.3sg.pres.subj money.gen.sg.f
 ‘whatever money each/any general has’ (Latin)⁶⁸

Zimmermann’s generalization states that every distance-distributive item that can be used as a determiner lacks the occasion reading. This is an implicational universal, not a biconditional. The opposite direction would be true if every distance-distributive item that lacks the occasion reading could be used as a determiner. Zimmermann considers Japanese as a counterexample to this opposite direction, but whether this is correct is not very clear. Zimmermann bases his view on the fact that the Japanese distance-distributive item *sorezore* differs formally from what he calls the Japanese distributive determiner-quantifier *wh . . . + mo*, which is illustrated in (47).

- (47) *Dono* *gakusei-mo* *sooseezi-o* *hutatu* *katta*.
 which student-MO sausage-acc two-cl bought
 ‘Every student bought two sausages.’ (Japanese)⁶⁹

However, *sorezore* can also be used in the position of a determiner, as example (34) shows. The syntactic status of *sorezore* in this example, and therefore the import of Japanese on Zimmermann’s generalization, is debatable, since Japanese is usually assumed to lack overt determiners. For a detailed semantic analysis of *sorezore*, see Sakaguchi (1998: ch. 3).

Setting *sorezore* aside, the inverse of Zimmermann’s generalization seems to hold for many languages. That is, when adnominal distance-distributive items cannot be used as determiners, they tend to have occasion readings. In addition to German *jeweils*, adnominal distance-distributive items that belong to this class are found in Bulgarian, Czech, Korean, Polish, Romanian, and Russian. Many of these observations are due to Zimmermann (2002b). The Korean case is discussed in depth by Choe (1987) and by Oh (2001, 2006), and the Polish case by Piñón (2000) and by Przepiórkowski (2013, 2014a, 2014b, 2015).

- (48) John i Mary kupiha *po* edna tetradka.
 John and Mary bought DIST one notebook
 ‘John and Mary bought one notebook each.’ (Bulgarian)⁷⁰

⁶⁷ Vergil, *Aeneid* 8, 660–661, Bortolussi (2013: 6). The slash stands for a line break in the *Aeneid*.

⁶⁸ Cicero, *De Lege Agraria* 1.10.8; see also Spevak (2014: 51).

⁶⁹ Satoshi Tomioka, p.c. to Zimmermann (2002b: 45).

⁷⁰ Petrova (2000), quoted in Zimmermann (2002b: 41).

- (49) *Vsjako/*Po* momče kupi dve nadenici.
 DIST boy bought two sausages
 'Each boy bought two sausages.' (Bulgarian)⁷¹
- (50) Mary byaga *po* 5 mili predi zakuska.
 Mary runs DIST 5 miles before breakfast
 'Mary runs five miles before breakfast (every morning).' (Bulgarian)⁷²
- (51) Chlapci koupili *po* dvou párcích/párkách.
 boys bought DIST two sausages.loc
 'The boys bought two sausages each.' (Czech)⁷³
- (52) *Každý/*Po* chlapec koupil dva párky.
 DIST boy bought two sausages
 'Each boy bought two sausages.' (Czech)⁷⁴
- (53) *Po* třech ženách vstupovalo do místnosti.
 DIST three.loc women.loc entered.3sg into room
 '[Each time,] three women entered the room.' (Czech)⁷⁵
- (54) Ai-tul-i phwungsen hana-ssik-ul sa-ess-ta.
 child-pl-nom balloon one-DIST-acc bought
 'The children bought one balloon each.' (Korean)⁷⁶
- (55) Sonyen-mata chayk-ul twu kwen-ssik sa-ess-ta.
 boy-DIST book-acc two cl-DIST bought
 'Every boy bought two books.' (Korean)⁷⁷
- (56) Na-nun phwungsen hana-ssik-ul sa-ess-ta.
 I-top balloon one-DIST-acc bought
 'I bought a balloon (each time / each day / at each store).' (Korean)⁷⁸
- (57) Chłopcy kupili *po* dwie kiełbaski.
 Boys bought DIST two sausages
 'The boys bought two sausages each.' (Polish)⁷⁹

⁷¹ Milena Petrova, p.c. to Zimmermann (2002b: 45); corrections supplied by Roumyana Pancheva, p.c. to the author.

⁷² Petrova (2000: ex. 3b), quoted in Zimmermann (2002b: 47).

⁷³ Hana Filip, p.c. to Zimmermann (2002b: 41).

⁷⁴ Hana Filip, p.c. to Zimmermann (2002b: 45) and to the author.

⁷⁵ Hana Filip, p.c. to Zimmermann (2002b: 47) and to the author.

⁷⁶ Choe (1987: 49, ex. 13) quoted in Zimmermann (2002b: 41); corrections supplied by WooJin Chung and Songhee Kim, p.c. to the author.

⁷⁷ Kim, p.c. to Zimmermann (2002b: 45); corrections supplied by WooJin Chung and Songhee Kim, p.c. to the author.

⁷⁸ Choe (1987: 52, ex. 18), quoted in Zimmermann (2002b: 47); corrections supplied by WooJin Chung and Songhee Kim, p.c. to the author.

⁷⁹ Adam Przepiórkowski, p.c. to the author.

- (58) *Każdy/*Po* chłopak kupił dwie kiełbaski.
 DIST boy bought two sausages
 'Each boy bought two sausages.' (Polish)⁸⁰
- (59) Papież zwiedzał *po* trzy kraje.
 Pope visited DIST three countries
 'The pope visited three countries each time.' (Polish)⁸¹
- (60) Doi oameni au cărat *cîte* trei valize.
 two men have carried DIST three suitcases
 'Two men have carried three suitcases each.' (Romanian)⁸²
- (61) *Fiecare/*Cîte* om cară trei valize.
 DIST man carry-pres.3sg three suitcases
 'Each man is carrying three suitcases.' (Romanian)⁸³
- (62) Un om cară *cîte* trei valize.
 one man carry-pres.3sg DIST three suitcases
 'One man carries three suitcases each time.' (Romanian)⁸⁴
- (63) Mal'chiki kupili *po* dve sosiski.
 boys buy.pfv.past.pl DIST two sausage
 'The boys bought two sausages each.' (Russian)⁸⁵
- (64) *Kazhdyj/*Po* mal'chik kupil dve sosiski.
 DIST boy buy.pfv.past.m.sg two sausage
 'Each boy bought two sausages.' (Russian)⁸⁶
- (65) Petja pokupal *po* dve sosiski.
 Petya buy.ipfv.past.m.sg DIST two sausage
 'Petya bought two sausages on each occasion.' (Russian)⁸⁷

Since Zimmermann's generalization is not about languages but about items, it can sometimes be observed within one language. I have illustrated this by including examples (42) through (44), which show that Russian *kazhdyj* can be used as a determiner and lacks the occasion reading, as well as examples (63) through (65), which show that Russian *po* has the occasion reading and cannot be used as a determiner.

⁸⁰ Adam Przepiórkowski, p.c. to the author.

⁸¹ Przepiórkowski (2014a: 110).

⁸² Gil (1982: 18, ex. 1f), Zimmermann (2002b: 41), Gil (1993: 298, ex. 66b).

⁸³ Gianina Iordăchioaia, p.c. to the author. See also Brasoveanu & Farkas (2011: 10).

⁸⁴ Gil (1993: 298, ex. 66a), Gianina Iordăchioaia, p.c. to the author. See also Iordăchioaia & Soare (2015: sect. 4).

⁸⁵ Masha Esipova, p.c. to the author.

⁸⁶ Masha Esipova, p.c. to the author. See also example (43).

⁸⁷ Masha Esipova, p.c. to the author.

Many languages express adnominal distance distributivity by a bound morpheme that attaches to a numeral—most commonly, a reduplicative morpheme (Gil 1982, 1993, 2013). The import of this fact for Zimmermann's generalization is unclear, since bound morphemes are not expected to be able to act as determiners, hence their inability to do so is not surprising. I mention it here for completeness and because the compositional analysis given in Section 9.3 extends to these morphemes. On the one hand, we find cases where reduplication does not give rise to occasion readings, such as Hungarian (Farkas 1997, Szabolcsi 2010):

- (66) A gyerekek két-két majmot láttak.
 The children two-two monkey.acc saw.3pl
 a. *Available*: 'Each of the children saw two monkeys.'
 b. *Unavailable*: 'The children saw two monkeys each time.' (Hungarian)⁸⁸

On the other hand, we find cases where bound morphemes do give rise to occasion readings, such as the reduplicative morphemes in Hausa (Zimmermann 2008), Karitiana (Müller & Negrão 2012), and Telugu (Balusu 2005, Balusu & Jayaseelan 2013), and the suffix in Tlingit (Cable 2014). All these cases are illustrated here and discussed in detail in their respective sources. For more examples and a typological overview, see Gil (1995, 2013).

- (67) yàaraa biyāṛ biyāṛ sun zoo.
 children five five 3pl.perf come
 a. 'The children came in groups of five.'
 b. 'On each occasion, five children came.' (Hausa)⁸⁹
- (68) Audù yaa sàyi lèemoo ukù ukù.
 Audu 3sg.perf buy orange three three
 'Audu bought oranges in threes.' (Hausa)⁹⁰
- (69) *Sypomp sypomp* nakam'at gooḡ ōwā.
 two-obl two-obl 3-decl-caus-build-nfut canoe child
 a. 'Each child built two canoes.'
 b. 'On each occasion, children built two canoes.' (Karitiana)⁹¹
- (70) pilla-lu *renDu renDu* kootu-lu-ni cuus-ee-ru.
 kid-pl two two monkey-pl-acc see-past-3pl
 a. 'The kids each saw two monkeys.'
 b. 'The kids saw two monkeys each time.'
 c. 'The kids saw monkeys in groups of two.' (Telugu)⁹²

⁸⁸ Szabolcsi (2010: 138, ex. 99).

⁸⁹ Zimmermann (2008: 462, ex. 98a).

⁹⁰ Zimmermann (2008: ex. 98b).

⁹¹ Müller & Negrão (2012: 160). The authors note that Karitiana bare nouns are number-neutral and not specified for definiteness.

⁹² Balusu & Jayaseelan (2013: 67, ex. 15a).

- (71) *Nás'gígáa* xáat has aawasháat.
 three.DIST fish pl.30.pfv.3s.catch
 a. 'They caught three fish each.'
 b. 'They caught three fish each time.' (Tlingit)⁹³

The facts discussed in this section suggest the following requirements for a semantic analysis of distance distributivity. First, the synonymy of the determiner, adnominal and adverbial uses of *each* in English should be captured, ideally by essentially identical lexical entries. Second, the fact that distance-distributive items across languages share some part of their meanings (namely their individual-distributive readings) should be represented, as well as the fact that some of them can also have occasion readings in suitable contexts. Third, the analysis should clarify the connections between distance-distributive items and distributivity theory more generally, and should capture the semantic variation across distance-distributive items. Finally, an explanation should be readily available for the crosslinguistic observation that distance-distributive items that can also be used as determiners can only distribute over individuals (Zimmermann's generalization). The rest of this chapter develops an analysis that fulfills these requirements.

The analysis I will propose is semantic, and does not aim to explain crosslinguistic differences that are syntactic in nature. For example, English-type languages are less likely to allow for the inverse distribution of subject over object denotations than some German-type languages are:

- (72) *One journalist each interviewed the politicians.
 (73) Jeweils ein Journalist interviewte die Politiker.
 DIST one journalist interviewed the politicians
 'The politicians were interviewed by one journalist each.' (German)

See Zimmermann (2002b: sects 5.4.2 and 5.4.3) for an extensive discussion and an integrated syntactic and semantic approach. My own account does not specify the syntactic constraints that govern the θ -indexing of distributivity operators. A separate theory of θ -indexing could be developed to account for the locality constraints on adnominal *each* constructions. These conditions have been studied in detail (e.g. Burzio 1986, Safir & Stowell 1988, Zimmermann 2002b: ch. 3). Other examples of coindexation that are subject to syntactic locality constraints are familiar from binding theory (Chomsky 1981, Büring 2005). The crosslinguistic variation on locality conditions is also discussed in Section 9.4.3.

9.3 Relating overt and covert distributivity

The connection between the D operator from Link (1987b) and adverbial *each* that was illustrated in (1) has been noted many times. As discussed in Section 3.4, there are

⁹³ Cable (2014: 564, ex. 3b).

different ways to implement this connection. We can treat *each* as synonymous with *for* and *of* except for the granularity and dimension parameter settings (see Chapter 4). In that case, *each* distributes indirectly, by presupposing that the verbal predicate is distributive. Alternatively, adverbial *each* can be treated as an overt realization of the D operator (Link 1991b). On this view, *each* distributes directly, by shifting the verbal predicate to a distributive interpretation. In this chapter, I adopt the second view, since my focus is on the connection between *each* and the D operator rather than between *each* and *for*-adverbials or pseudopartitives.

Here, I will stress the connection between *each* and D operators. Accordingly, I take adverbial and adnominal *each* and related distance-distributive items in Albanian, Dutch, French, Icelandic, Italian, Japanese, Norwegian, Portuguese, Russian, and possibly Latin to be essentially D operators. These are the languages mentioned in Section 9.2 as being English-type. As for *jeweils* and its relatives in German-type languages like Bulgarian, Czech, Korean, Polish, Romanian, and Russian, we have seen that they can distribute over spatial and temporal intervals—arguably nonatomic entities. Link's D operator always distributes down to individual atoms and can therefore not be extended to these cases. I will connect them to the nonatomic distributivity operator Part from Schwarzschild (1996). For the sake of brevity, I will only execute the analysis for English *each* and German *jeweils*, but it should be clear how to extend it to other distributive items depending on which one of these two they pattern with. I will only show how to model the individual-distributive and the temporal occasion readings. The extension from the temporal to the spatial occasion reading is straightforward.

The guiding idea of the analysis is that overt distributive items include two versions of the distributivity operator. *Each* includes the atomic distributivity operator D, which can distribute only over count domains because only those domains have atoms. *Jeweils* includes the nonatomic distributivity operator Part. I have argued in Section 8.6 that the latter operator can also distribute over noncount domains like time. As in the rest of this book, I adopt the strata-theoretic perspective, according to which distributivity is a property with two parameters: dimension and granularity. I will suggest that *each*, just like the D operator, comes prespecified for "granularity=atom." This blocks the setting "dimension=time," hence distributivity over occasions is unavailable. By contrast, *jeweils* does not come prespecified for anything but is anaphoric on the context. It can therefore distribute over salient covers, or salient stretches of time, just like the Part operator.

In claiming that *each* is an overt form of the D operator, I loosely follow proposals made for German *jeweils* and its short form *je* by Link (1987b, 1998b) and for English *each* by Roberts (1987). While Link and Roberts did not give explicit compositional implementations and did not fully consider the crosslinguistic picture, this chapter can be seen as an update to their ideas which benefits from later work on algebraic semantics, nonatomic distributivity, and compositional implementations.

Arriving at the meaning of *each* from the meaning of distributivity operators is somewhat the reverse of the process by which Schwarzschild arrived at his Part operator, which “was based on a generalization of Dowty & Brodie’s (1984) account of floated quantifiers as verb phrase modifiers” (Schwarzschild 1996: 137). Schwarzschild himself notes that the history of Part should not be taken for an endorsement that floated quantifiers are related to it, and argues that floated quantifiers should be distinguished from distributivity operators because he takes reciprocals to be licensed by distributivity operators, but not by adverbial *each*. He gives the examples in (74) to support this claim. These kinds of examples, as well as the idea that reciprocals are licensed by distributivity operators of some kind or other, go back to Heim, Lasnik & May (1991).

- (74) a. They_j Part_i [saw each other_{j,i}].
 b. *They_j each_i saw each other_{j,i}.

Schwarzschild’s argument rests on the assumption that reciprocals are licensed by VP-level distributivity operators. But then we should expect that all VPs with reciprocals in them are interpreted distributively, contrary to fact (see Dotlačil 2013 and references therein):

- (75) a. John and Mary wrote to each other on two cold days.
 ?? under the reading ‘John wrote to Mary on two cold days and Mary wrote to John on two other cold days’ (Moltmann 1992)
 b. The doctors gave each other a new nose.
 ?? under the reading ‘each doctor gave the other doctor a different new nose’ (Williams 1991)
 c. The two children gave each other a Christmas present.
 ?? under the reading ‘each child giving a different present’ (Williams 1991)

Since neither overt nor covert adverbial distributive operators seem to be able to license reciprocals, nothing stands in the way of a unified analysis, to which I turn now.

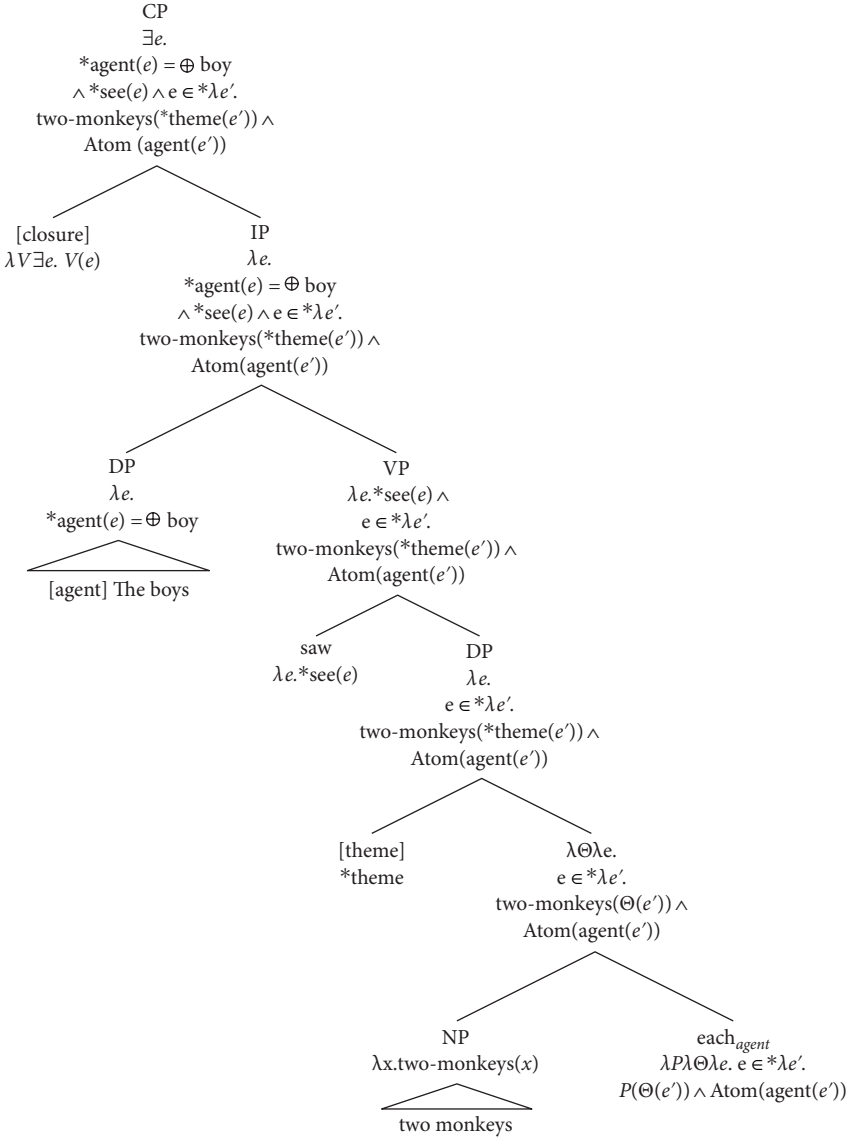
Here are the entries for adverbial and adnominal *each*; determiner *each* is discussed in Section 9.5. An explanation follows. I assume that adverbial *each*, as shown in (76), is a verb phrase modifier just like the D operator, and can therefore be given the same entry as that operator, repeated here from Section 8.4. An illustration of the derivation of a basic sentence with adnominal *each* is shown in Figure 9.1. Adverbial *each* works similarly.

$$(76) \llbracket \text{each}_\theta \rrbracket_{\text{adverbial}} = \llbracket D_\theta \rrbracket = (78)$$

$$(77) \llbracket \text{each}_\theta \rrbracket_{\text{adnominal}} = \lambda P \lambda \Theta \lambda e [e \in \llbracket D_\theta \rrbracket (\lambda e'. P(\Theta(e')))]$$

$$(78) \text{ Event-based D operator (repeated from Section 8.4)}$$

$$\llbracket D_\theta \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in {}^* \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$$

FIG. 9.1 Deriving *The boys saw two monkeys each*.

I adopt for concreteness the assumption that adverbial *each* is an adverb adjoined to VP. This is similar to what has been argued for floating quantifiers in general by Dowty & Brodie (1984), Bobaljik (1995), and Doetjes (1997). Another view analyzes floating quantifiers as the remaining part of a noun phrase the rest of which has moved away from it (e.g. Safir & Stowell 1988, Sportiche 1988). The movement view makes a

formal link between *each* and its antecedent available for independent reasons since there is a movement relation between them. The adverbial view leaves it open whether a formal link is created (e.g. via θ -indexing, as I assume here for English *each*) or whether the target of *each* is determined in some other way, for example by choosing from a small inventory of thematic roles and similar functions. In Section 9.4.3, I consider this possibility for German *jeweils*.

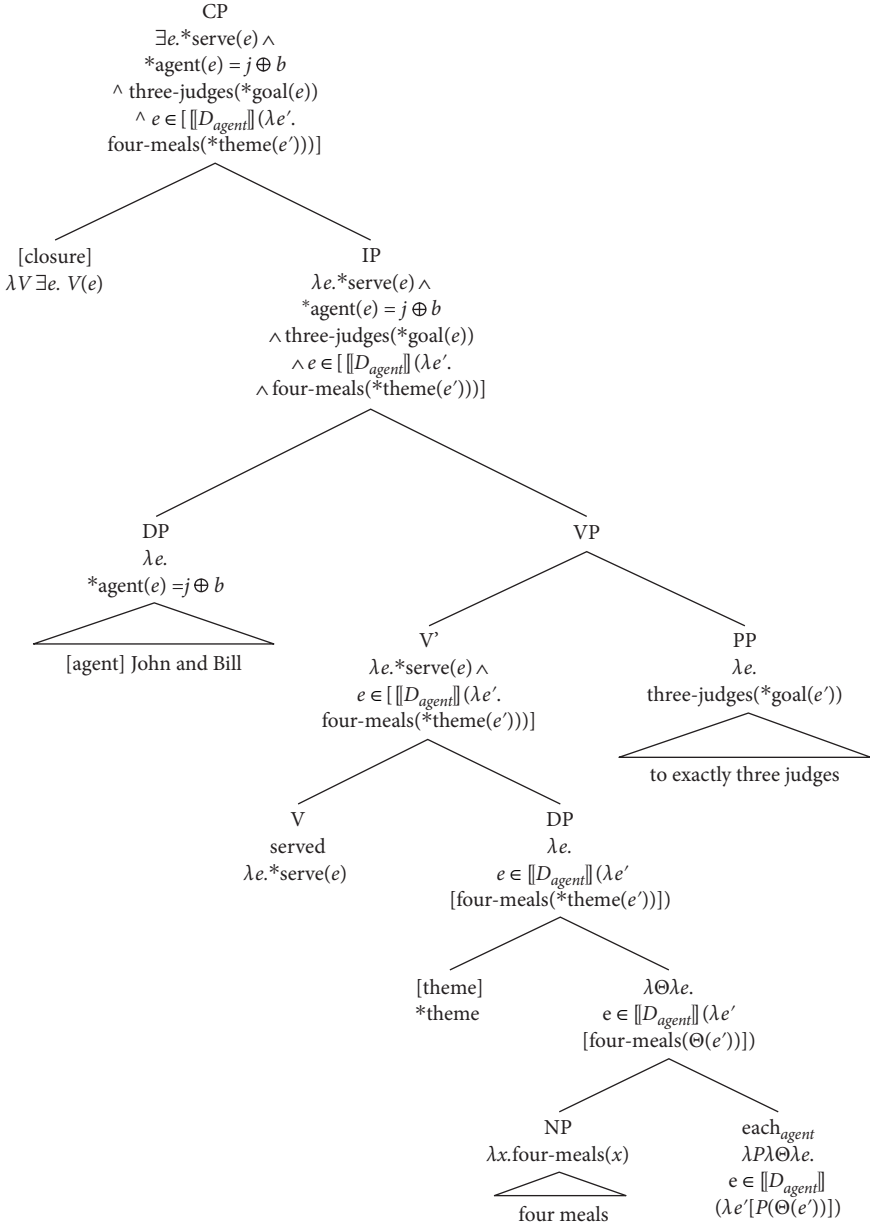
On the present view, adverbial *each* is synonymous with the D operator, and adnominal *each* is essentially a type-shifted D operator. This captures the fact that they are essentially synonymous to each other. As shown in (77), adnominal *each* carries an index, which I assume is θ -indexed with the thematic role of its antecedent, written as θ . In the compositional derivation, adnominal *each* first combines with its host predicate *P* (e.g. *two monkeys*), and then with the θ -head of its host, written as Θ (not to be confused with the θ -role of its antecedent). Afterwards, it combines intersectively with the verbal projection to which its host attaches (e.g. the verb *see*). This means that adnominal *each* takes scope over its complement but—unlike adverbial *each*—not over the verbal projection (Dotlačil 2011, 2012, LaTerza 2014a, 2014b). A previous version of my theory, Champollion (2012), gave adnominal *each* scope over the verb phrase as well. This leads to wrong predictions as discussed in LaTerza (2014a, 2014b), and has been fixed here. The problem can be illustrated with the minimal pair in (79) (LaTerza 2014b):

- (79) a. John and Bill served [[four meals] each] to (exactly) three judges.
 b. John and Bill [each [served four meals to (exactly) three judges]].

As LaTerza reports, speakers judge (79a) true of situations where there are at most three judges, while (79b) is true in situations which allow up to six different judges. This suggests that adnominal and adverbial *each* take scope as indicated by the square brackets in these examples. Sentence (79a) can be derived as in Figure 9.2.

My entry for adnominal *each* combines with its host in two steps, in order to give it access to both the predicate and the θ -head. This is not essential, but it allows us to ensure that the type of the predicate is $\langle e, t \rangle$. I do so to provide a hook on which to build future accounts of the “counting quantifier requirement” that prevents such phrases as **most men each* (Safir & Stowell 1988, Sutton 1993, Szabolcsi 2010: sect. 10.5). The theory in this chapter does not aim to provide an account of this requirement and will not rule out bare plurals as in **They saw monkeys each*, as pointed out in Cable (2014). If an independent account of these kinds of mismatches can be provided that does not need separate access to the host predicate and its θ -role, it may not be necessary to place *each* between the host predicate and the θ -head after all.

Turning now to *jeweils*, my reformulation of the Part operator in Section 8.4, repeated here as (80), provides the basis for its lexical entries.

FIG. 9.2 Deriving *John and Bill served four meals each to exactly three judges*.

(80) **Event-based Part operator** (repeated from Section 8.4)

$$\llbracket \text{Part}_{\theta, C} \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right)$$

(Takes an event predicate V and returns a predicate that holds of any event e which can be divided into events that are in V and whose θ s satisfy the contextually salient predicate C .)

Adverbial *jewels* is treated as in (81).

$$\begin{aligned} (81) \quad & \llbracket \text{jewels}_{\theta, C} \rrbracket_{\text{adverbial}} \\ &= \lambda V \lambda e [\llbracket \text{Part}_{\theta, C} \rrbracket(V)(e) \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e))] \\ &= \lambda V \lambda e \left[e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right) \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e)) \right] \end{aligned}$$

As in the case of *each*, I assume that the free variable θ can be resolved through coindexation with a thematic relation. Unlike *each*, however, this is not required. In other cases to be discussed shortly, I will assume that it can also be resolved to other values such as τ (runtime) or *id* (the identity function). As for the free variable C , I assume that pragmatics ensures that it is either resolved to *Atom* or to a set whose elements are contextually familiar (see Schwarzschild 1996). In the latter case, the second conjunct of (81) ensures that C is an exact cover of $\theta(e)$, in the sense that the members of C sum up exactly to $\theta(e)$ as opposed to some entity that properly contains $\theta(e)$; the point of this requirement will become clear at the end of this section. In the case of *each*, it was not necessary to state it because *each* only distributes over atoms in the first place. Atoms are stipulated to be exempt from this requirement because the set of atoms covers most of the domain of discourse.

The same type shift as in (77), modulo the exact-cover requirement, brings us from (81) to adnominal *jewels*:

$$\begin{aligned} (82) \quad & \llbracket \text{jewels}_{\theta, C} \rrbracket_{\text{adnominal}} \\ &= \lambda P \lambda \Theta \lambda e [\llbracket \text{Part}_{\theta, C} \rrbracket(\lambda e' [P(\Theta(e'))])(e) \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e))] \\ &= \lambda P \lambda \Theta \lambda e \left[e \in * \lambda e' \left(\begin{array}{c} P(\Theta(e')) \wedge \\ C(\theta(e')) \end{array} \right) \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = \theta(e)) \right] \end{aligned}$$

In Section 9.4.4, we will encounter another instance of adnominal *jewels*, whose meaning is identical to adverbial *jewels* except that it distributes over individuals rather than events. One could easily unify it with (81) by treating its variables as untyped or type-polymorphic. More generally, it should be obvious that all these lexical entries have the same semantic common core. Still, they differ in their types and in the number and types of their arguments because they are formulated in a way that allows them to be used in syntax trees that correspond closely to the surface forms of sentences. This seems to me to be a reasonable tradeoff. For a related account, which makes the lexical entries of various distance-distributive items even more similar at the cost of adding more empty elements to the syntax, see LaTerza (2014a).

As in the case of the Part operator, the granularity parameter C of *jeweils* can be set to *Atom* so long as its dimension parameter θ is set to a function into a count domain, such as *agent*. In that case, Part distributes over individuals and is equivalent to the D operator, as explained in Section 8.4. This accounts for the fact that when *jeweils* distributes over individuals, it is equivalent to *each*, as this German sentence illustrates:

- (83) Die Jungen haben *jeweils*_{agent, Atom} zwei Affen gesehen.
 The boys have DIST two monkeys seen
 ‘The boys have each seen two monkeys.’ = (9)

If—and only if—there is a supporting context, the anaphoric predicate C can be set to a salient antecedent other than *Atom*. In that case, θ is free to adopt values with nonatomic ranges, such as τ (runtime). This leads to occasion readings. Suppose for example that (83) is uttered in a context where it is in the common ground that the boys have been to the zoo three times recently. The predicate that contains these three time intervals, call it *zoovisit*, is salient in this context. The derivation proceeds along similar lines to what is shown in Figure 9.1 and yields the result shown in (84):

- (84) [[Die Jungen haben *jeweils* _{$\tau, zoovisit$} zwei Affen gesehen.]] =
 $\exists e. *agent(e) = \bigoplus \text{boy} \wedge *see(e)$
 $\wedge e \in *\lambda e' \left(\begin{array}{l} |*theme(e')| = 2 \wedge *monkey(*theme(e')) \wedge \\ zoovisit(\tau(e')) \end{array} \right)$
 $\wedge \bigoplus zoovisit = \tau(e)$
 ‘The boys have seen two monkeys on each salient occasion (i.e. on each of the three zoo visits).’

Since *zoovisit* is salient, C can be resolved to it rather than to *Atom*. Since there are no atoms in time, it is only now that θ can be set to τ , rather than to *agent* as in (83). What (84) asserts in this context is that there is an event e whose (discontinuous) runtime is the sum of the three zoo visits; that this event has the boys as its agents; that it can be divided into subevents, each of whose runtimes is the time of a zoo visit; and that each of these subevents is an event whose theme is two monkeys. That these subevents are seeing events is entailed by the fact that *see* is lexically distributive on its theme argument, which in turn can be formally represented as a meaning postulate, as discussed in Section 4.5. I assume that runtime is closed under sum just like other thematic roles ($\tau = *\tau$), or in other words, it is a sum homomorphism (see Section 2.5.2). This means that any way of dividing e must result in parts whose runtimes sum up to $\tau(e)$. The conjunct $\bigoplus zoovisit = \tau(e)$ makes sure that $\tau(e)$ is the sum of the times of the three zoo visits in question. Hence each of these zoo visits must be the runtime of one of the seeing-two-monkeys events. This improves on an earlier version of this theory which lacked the conjunct in question (Champollion 2012).

Without it, (84) would be predicted true even if the boys failed to see two monkeys on some of the salient zoo visits.

9.4 Some more complicated cases

To demonstrate the viability and versatility of the present analysis, I will now apply it to a few configurations that are more complicated than those discussed so far. The subsequent sections do not depend on this section. Detailed syntactic and semantic analyses of many of the configurations discussed here (and many others) are found in Zimmermann (2002b). Section 9.7 is devoted to a critical review of the semantic aspects of that account.

9.4.1 *Each as a PP modifier*

Each can occur as the modifier of a prepositional phrase. Example (85) is a simple case. Example (86) plays an important role in Schein (1993) and has not received a compositional semantic analysis so far.

(85) Mary put the books each back on the bookshelf. (Maling 1976)

(86) 300 quilt patches covered two workbenches each with two bedspreads. (Schein 1993)

To analyze these sentences, I assume that *each* modifies the prepositional phrase to its right, similarly to the adverb *back* in *back at the farm*, rather than the noun phrase to its left. (As (85) shows, these modifiers can be stacked.) My assumption is plausible because adnominal *each* cannot modify definite plurals like *the books*. I assume for concreteness that the syntactic structure of these sentences is [[V DP] PP] rather than [V [DP PP]]; for discussion on the choice between these two analyses, see Janke & Neeleman (2012).

Example (86) has a reading according to which there are a total of two workbenches and a total of four bedspreads that cover them. The workbenches stand in a cumulative relation with the 300 quilt patches. The following formula captures this reading:

- (87) $\exists e. *cover(e)$
 $\wedge *quilt-patch(*theme(e)) \wedge |*theme(e)| = 300$
 $\wedge *workbench(*goal(e)) \wedge |*goal(e)| = 2$
 $\wedge e \in * \lambda e' \left(\begin{array}{l} *bedspread(*instrument(e')) \wedge |*instrument(e')| = 2 \\ \wedge Atom(goal(e')) \end{array} \right)$
 (There is a sum of covering events whose themes sum up to 300 quilt patches, whose goals sum up to two workbenches, and which can be divided into two smaller sum events, each of which involves two bedspreads and one of the workbenches.)

Formula (87) is derived as follows. I have used shortcuts like *300-quilt-patches* for better readability. The derivation is straightforward and does not make use of any new ingredients. The entry for *each* is the same as the adverbial one, even though it modifies a prepositional phrase and not a verb phrase. This works because the prepositional phrase is represented as an event predicate, just like a verb phrase.

$$(88) \llbracket \text{each}_{\text{goal}} \rrbracket = \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\text{goal}(e')) \end{array} \right)$$

$$(89) \llbracket \text{with two bedspreads} \rrbracket = \lambda e. \text{two-bedspreads}(*\text{instrument}(e))$$

$$(90) \llbracket \text{each}_{\text{goal}} \text{ with two bedspreads} \rrbracket \\ = \lambda e. e \in * \lambda e' \left(\begin{array}{c} \text{two-bedspreads}(*\text{instrument}(e')) \wedge \\ \text{Atom}(\text{goal}(e')) \end{array} \right)$$

$$(91) \llbracket [\text{goal}] \text{ covered two workbenches} \rrbracket \\ = \lambda e. * \text{cover}(e) \wedge \text{two-workbenches}(*\text{goal}(e))$$

$$(92) \llbracket [\text{theme}] \text{ 300 quilt patches} \rrbracket = \lambda e. \text{300-quilt-patches}(*\text{theme}(e))$$

$$(93) \llbracket (86) \rrbracket = \exists e. e \in (92) \cap (91) \cap (90) = (87)$$

9.4.2 *Jewels distributing over a conjunction of verbs*

German *jewels* can take a conjunction of verbs as its antecedent and distribute over the two events described by the conjuncts Moltmann (1997: 207). This results in a meaning for which English uses the word *respectively*:

- (94) Peter kritisierte und lobte Maria aus jeweils zwei Gründen.
 Peter criticized and praised Mary for DIST two reasons
 ‘Peter criticized and praised Mary for two reasons respectively.’
 (Zimmermann 2002b: 46)

As for English *each*, it cannot be used for that purpose:

- (95) *Peter criticized and praised Mary for two reasons each.
 (Zimmermann 2002b: 134)

According to Zimmermann (2002b: 143f.), other languages that pattern with English in this respect include Bulgarian, Czech, Dutch, French, Italian, Norwegian, and Russian. As we have seen in Section 9.2, this list includes many languages with distance-distributive items that can also be used as determiners and lack the occasion reading. I have suggested earlier that the occasion reading is only possible if the granularity parameter can be set to a nonatomic value. Therefore, distributivity over conjuncts is predicted to be impossible as long as the events described by the two conjuncts are nonatomic (pace Zimmermann 2002a). There is ample reason to assume

that they are indeed nonatomic. For one thing, *praise* and *criticize* are atelic predicates, so any praising event that goes on for five minutes will have parts that take less than five minutes. As another example, sentence (96) (suggested by a reviewer) entails that each of the six students was either praised or criticized, which means that the two verbs are lexically distributive on their themes (see Section 4.5). This in turn means that the praising event in (96) consists of three praising subevents, and similarly for the criticizing event.

- (96) Der Professor hat jeweils drei Studenten gelobt und kritisiert.
 The professor has DIST three students praised and criticized
 ‘The professor praised three students and criticized three students.’
 (German)

This explanation will work for most of the languages just mentioned, but not for all of them. As we have seen in Section 9.2, in Bulgarian and Czech the distance-distributive item *po* can be used to distribute over salient occasions and cannot be used as a distributive determiner. We would therefore expect that this item allows distribution over conjuncts, but it does not. Like Zimmermann, I have no explanation for this fact.

To derive (94) compositionally, I assume that the dimension parameter θ of *jeweils* is resolved to the identity function *id* rather than to a thematic role. As *jeweils* is not syntactically required to be coindexed with a thematic role, it is natural to assume that there are other salient functions that are licit values to be picked up by its dimension parameter. I also assume that the cover variable *C* is resolved to the pragmatically salient cover $\{e_c, e_p\}$ where e_c is the salient criticizing-for-two-reasons event and e_p is the salient praising-for-two-reasons event.

$$(97) \llbracket \text{jeweils}_{id,C} \rrbracket_{adnominal} \\ = \lambda P \lambda \Theta \lambda e. e \in * \left(\frac{P(\Theta(e')) \wedge}{C(e')} \right) \wedge (C \neq \text{Atom} \rightarrow \bigoplus C = e)$$

$$(98) \llbracket \text{jeweils}_{id,\{e_c,e_p\}} \text{zwei Gründen} \rrbracket \\ = \lambda \Theta \lambda e. e \in * \lambda e' \left(\frac{2\text{-reasons}(\Theta(e')) \wedge}{e' \in \{e_c, e_p\}} \right) \wedge e_c \oplus e_p = e$$

I assume that *aus* (in this context) denotes a function from events to their causes (or whatever is the relation between a praising/criticizing event and its reason).

$$(99) \llbracket \text{aus} \rrbracket = \lambda e. * \text{cause}(e)$$

$$(100) \llbracket \text{aus jeweils}_{id,\{e_c,e_p\}} \text{zwei Gründen} \rrbracket \\ = \lambda e. e \in * \lambda e' \left(\frac{2\text{-reasons}(* \text{cause}(e')) \wedge}{e' \in \{e_c, e_p\}} \right) \wedge e_c \oplus e_p = e$$

I represent the denotation of the verbal conjunction using sum formation as in (101). I remain noncommittal about the compositional derivation of this conjunction. For present purposes, we do not need to choose between a sum-based denotation of *and*, as in Lasersohn (1995), and an intersective denotation that involves Montague-lifting the two event predicates and then minimizing their intersection, as in Winter (2001) and Champollion (2015d, 2016d).

- (101) $\llbracket \text{kritisierte und lobte} \rrbracket$
 $= \lambda e \exists e_1 \exists e_2. * \text{criticize}(e_1) \wedge * \text{praise}(e_2) \wedge e = e_1 \oplus e_2$

Once all these building blocks have been put together and conjoined with the agent and theme, the result is as follows:

- (102) $\llbracket (94) \rrbracket = \exists e. \text{agent}(e) = \text{peter} \wedge \text{theme}(e) = \text{maria} \wedge$
 $\exists e_1 \exists e_2. * \text{criticize}(e_1) \wedge * \text{praise}(e_2) \wedge e = e_1 \oplus e_2 \wedge$
 $e \in * \lambda e' \left(\begin{array}{c} 2\text{-reasons}(* \text{cause}(e')) \wedge \\ e' \in \{e_c, e_p\} \end{array} \right) \wedge e_c \oplus e_p = e$

This is true just in case there is an event whose agent is Peter, whose theme is Maria, and which consists of two subevents e_c and e_p such that one of them is a criticizing event, the other one is a praising event, and each of these subevents is caused by two reasons. By thematic uniqueness, each of these two events has Peter as agent and Maria as theme.

9.4.3 *Jeweils in subject position*

As we have already seen in (8), German adnominal *jeweils* can occur as part of the subject of a clause (Zimmermann 2002b: 27):

- (103) *Jeweils ein Offizier begleitete die Ballerinen nach Hause.*
 DIST one officer accompanied the ballerinas to home
 ‘The ballerinas were accompanied home by one officer each.’
 (Zimmermann 2002b)

When the subject is at the beginning of the clause, as in (8) and (103), one may speak of “backwards distributivity” since the antecedent of *jeweils* occurs to its right. In English, backwards distributivity appears to be restricted to passives (Burzio 1986, Safir & Stowell 1988) and unaccusatives:

- (104) *One officer each accompanied the ballerinas home. (Zimmermann 2002b)
 (105) ?One interpreter each was assigned to the visitors. (Burzio 1986: 200)
 (106) Table 3 shows the dissertation topics for those holding earned doctorates. [...] Three dissertations each dealt with assessment, transfer, trustees, and technical education. Two dissertations each were on accreditation, counseling,

effectiveness, and mission. One dissertation each focused on economic development, learning resources, performing arts, and strategic planning.⁹⁴

- (107) Indeed, Mr. Mitsotakis commanded only 144 seats [...] The Socialists won 125 seats [...] and one seat each went to a conservative independent and to an ethnic Turk from Thrace, near the Turkish border.⁹⁵

I do not have a semantic explanation for the restriction against backwards distributivity in English. As in the case of locality constraints, I assume that this restriction can be dealt with by syntactic accounts such as the ones already proposed, for example by Safir & Stowell (1988). As one reviewer suggests, one might expect θ -indexing to turn out to obey a hierarchy comparable to the hierarchies of thematic roles that are sometimes claimed to be at work in binding theory (Jackendoff 1972: 148; see also Buring 2005: 16). If correct, this may help explain why attested cases of inverse distribution in English, such as the ones we have seen in (106) and (107), tend to involve nonagent subjects.

For the German case, where there is no restriction, my account can easily be used to derive the meaning of (103) if we assume that the dimension parameter of *jeweils* is provided by the thematic role of *die Ballerinen*. Here are the core elements of the derivation; I assume that *ein Offizier* is interpreted predicatively. Here and below, I omit the exact-cover conjunct ($C \neq \text{Atom} \rightarrow \bigoplus C = e$) whenever it is vacuously true.

- (108) $\llbracket \text{jeweils}_{\text{theme}, \text{Atom}} \rrbracket$ (adnominal)
 $= \lambda P \lambda \Theta \lambda e. e \in * \lambda e' \left(\begin{array}{c} P(\Theta(e')) \wedge \\ \text{Atom}(\text{theme}(e')) \end{array} \right)$
- (109) $\llbracket [\text{agent}] \text{jeweils}_{\text{theme}, \text{Atom}} \text{ein Offizier} \rrbracket$
 $= \lambda e. e \in * \lambda e' \left(\begin{array}{c} \text{officer}(\text{agent}(e')) \wedge \\ \text{Atom}(\text{theme}(e')) \end{array} \right)$
- (110) $\llbracket \text{begleitet} [\text{theme}] \text{die Ballerinen} \rrbracket$
 $= \lambda e. * \text{accompany}(e) \wedge * \text{theme}(e) = \bigoplus \text{ballerina}$
- (111) $\llbracket [\text{agent}] \text{jeweils}_{\text{theme}, \text{Atom}} \text{ein Offizier begleitet} [\text{theme}] \text{die Ballerinen} \rrbracket$
 $\exists e. * \text{accompany}(e) \wedge * \text{theme}(e) = \bigoplus \text{ballerina} \wedge$
 $e \in * \lambda e' \left(\begin{array}{c} \text{officer}(\text{agent}(e')) \wedge \\ \text{Atom}(\text{theme}(e')) \end{array} \right)$

What (111) says is that there is a sum of accompanying events whose themes sum up to the ballerinas and which consists of parts e' such that each e' has an atomic theme and an officer as its agent.

⁹⁴ Attested example, Keim & Murray (2008: 125f.).

⁹⁵ Attested example, *New York Times*, "Greek Conservative Is Seeking Coalition" (June 20, 1989).

These are the correct truth conditions for the German sentence (103). But given that its English counterpart (104) is ruled out, why is (103) acceptable? This question needs to be answered by a syntactic theory, and I can only offer speculation. One possible explanation is that there are language-specific constraints on θ -indexing. This is consistent with the fact that even in German, there appears to be a clausemate requirement between *jeweils* and its antecedent (Zimmermann 2002b: 26f.). Another possibility is that distance-distributive items differ across languages in whether they require θ -indexing in order to distribute over another element in the sentence.⁹⁶ On this view, *jeweils* in (103) is not actually coindexed with the θ -role of the ballerinas. Rather, it distributes over a set of occasions which stand in a one-to-one relation with the ballerinas and which are made salient by the fact that the ballerinas are mentioned in the sentence. If this is correct, there is no formal link between *jeweils* and the ballerinas. This makes an interesting prediction: The languages that allow distance-distributive items in subject position should be just the ones that allow distribution over salient entities that need not be overtly mentioned and need not be atomic. This prediction indeed appears to be borne out (Zimmermann 2002b: 48–50): Aside from German, at least Bulgarian, Czech, Korean, and Polish have distance-distributive items that can occur in subject position (for Polish, see Przepiórkowski 2013). Aside from English, at least Dutch, French, Icelandic, Italian, Norwegian, and Russian have distance-distributive items that cannot occur in subject position (setting aside English passives and unaccusatives). We saw in Section 9.2 that distance-distributive items in the first set of languages allow distribution over salient nonatomic entities, while those in the second set do not.

9.4.4 Reverse DP-internal distributivity

The analysis can be extended to a configuration halfway between the adverbial and adnominal case: backwards distributivity within a noun phrase. I illustrate this case with a German example:

- (112) Das Parlament hat jeweils zwei Abgeordnete aus den drei
 The parliament has DIST two representatives from the.acc.pl three
 baltischen Staaten eingeladen.
 Baltic states invited
 ‘From each of the three Baltic states, two representatives were invited by the
 parliament.’ (German)

A parallel construction is available with Polish distributive *po*; both in German and in Polish, this configuration poses a problem for the account in Zimmermann (2002b), as discussed in detail by Przepiórkowski (2014a, 2015). As I show here, the

⁹⁶ The clausemate requirement must then be due to something else than a formal link between *jeweils* and its antecedent. For example, it could be due to a requirement that they modify the same event.

present account can be extended straightforwardly to this kind of configuration. For a semantic analysis in a different framework, see also Przepiórkowski (2014a, 2014b).

I write $es \oplus la \oplus li$ for the sum of the three Baltic states, Estonia, Latvia, and Lithuania. I assume that in this context, *aus* ('from') denotes a function that maps individuals to their origins, and that is closed under sum:

$$(113) \quad \llbracket \text{aus} \rrbracket = \lambda y \lambda x. *origin(x) = y$$

The prepositional phrase then denotes the set of all plural individuals whose origins are the three Baltic states:

$$(114) \quad \llbracket \text{aus den drei baltischen Staaten} \rrbracket = \lambda x. *origin(x) = es \oplus la \oplus li$$

The complex noun phrase denotes a sum of six representatives consisting of three pairs, with each pair coming from one of the three Baltic states. Although this instance of *jeweils* is adnominal, it has the denotation of adverbial *jeweils* in (8o) except that it ranges over individuals instead of events. In the sentence at hand, its dimension parameter is set to the *origin* function I used as the denotation of *aus*, and its granularity parameter to *Atom* (since each pair of representatives comes from a single Baltic state).

$$(115) \quad \llbracket \text{jeweils}_{origin, Atom} \rrbracket = \lambda P \lambda x. x \in * \lambda y \left(\begin{array}{c} P(y) \wedge \\ Atom(origin(y)) \end{array} \right)$$

$$(116) \quad \llbracket \text{jeweils}_{origin, Atom} \text{ zwei Abgeordnete} \rrbracket \\ = \lambda x. x \in * \lambda y \left(\begin{array}{c} |y| = 2 \wedge *representative(y) \wedge \\ Atom(origin(y)) \end{array} \right)$$

The denotation of the complex noun phrase is the intersection of (114) and (116). After it combines with the *theme* θ -head via the appropriate type shifter (see Section 2.10), the result is this:

$$(117) \quad \llbracket [\text{theme}] \text{jeweils zwei Abgeordnete aus den drei baltischen Staaten} \rrbracket \\ = \lambda e. *origin(*theme(e)) = es \oplus la \oplus li \wedge \\ *theme(e) \in * \lambda y \left(\begin{array}{c} |y| = 2 \wedge *representative(y) \wedge \\ Atom(origin(y)) \end{array} \right)$$

After combining with the main verb *eingeladen* and with the subject *Das Parlament*, the final result is as follows:

$$(118) \quad \llbracket (112) \rrbracket \\ = \exists e. *agent(e) = ix[\text{parliament}(x)] \wedge \\ *invite(e) \wedge *origin(*theme(e)) = es \oplus la \oplus li \wedge \\ *theme(e) \in * \lambda y \left(\begin{array}{c} |y| = 2 \wedge *representative(y) \wedge \\ Atom(origin(y)) \end{array} \right)$$

This says that there is an inviting event whose agent is the parliament, and whose theme has the following properties: Its origins sum up to the three Baltic states, and it consists of sums of two representatives, each of which has a single country as its origin. These are the desired truth conditions.

The present analysis could no doubt be improved, for example by generalizing the dimension parameter from functions to relations so that its values are not restricted to thematic roles, functions like *runtime*, and function-denoting prepositions like *aus* ‘from’. This would further increase its empirical coverage. I have not done so because my goals here do not include accounting for every possible syntactic configuration in which distance-distributive items can be used.

While the analysis so far has focused on adnominal and adverbial *each* and their counterparts across languages, it is possible to assimilate distributive determiners such as *each* and *every* to these items. I turn to them now.

9.5 Distributive determiners

As we have seen, English *each* along with some of its crosslinguistic relatives can be used adnominally, adverbially, and as a determiner. I have suggested that the synonymy of these uses should be captured, ideally by essentially identical lexical entries. Another distributive determiner in English is *every*. As shown by their incompatibility with collective predicates, both *every* and *each* are distributive (see Section 4.2.1):

- (119) #Every/#Each soldier surrounded the castle. (Kroch 1974: ch. 5)

Traditionally, the determiners *every* and *each* are analyzed in terms of universal quantification (e.g. Montague 1973):

- (120) $\llbracket \text{every boy} \rrbracket = \lambda P \forall x [\text{boy}(x) \rightarrow P(x)]$ (traditional)

This style of analysis is especially useful when one is interested in comparing them with other determiners from the perspective of generalized quantifier theory (e.g. Barwise & Cooper 1981). This chapter, however, focuses on the parallels between determiner *each* and its adnominal and adverbial counterparts. Therefore, instead of the traditional approach I will reuse the analyses of adverbial and adnominal *each* that we have already encountered. Since the differences between *each* and *every* are not a core concern of this chapter, I will adopt the same analysis for both determiners. This is not to deny that there are differences between them. To mention some examples, determiner *each*, unlike *every*, is not clause-bounded (Szabolcsi 2010: 107). It has a strong preference for taking wide scope over its environment, more so than *every* (e.g. Ioup 1975, Beghelli 1997, Beghelli & Stowell 1997, Tunstall 1998). Relatedly, *each* appears to impose a differentiation requirement on its subevents (Tunstall 1998, Brasoveanu & Dotlačil 2015, Thomas & Sudo 2016):

- (121) a. A helper dyed every shirt. *both scopal orders possible*
 b. A helper dyed each shirt. *inverse scope strongly preferred*
- (122) Jake photographed { every / #each } student in the class, but not individually.

Another difference is that *each* can but *every* cannot readily be used to quantify over a set having only two members:

- (123) a. { #Every one / Each } of the two ... (Vendler 1967: 77)
 b. Cromwell held a bible in { #every / each } hand. (Aldridge 1982: 218)

Conversely, a variety of environments tolerate *every* but not *each*:

- (124) a. { Every / ?Each } ten weeks, he visits Spain. (Aldridge 1982: 221)
 b. It took { every / *each } boy to lift the piano. (Beghelli & Stowell 1997: 98)
 c. Not { every / *each } boy ate an ice-cream cone. (Beghelli & Stowell 1997: 99)
 d. Almost { every / *each } student left the room. (Farkas 1997: sect. 3)

Discussions of these and further differences between *every* and *each* can be found in Vendler (1967: 76–9), Hogg (1977: 135–9), Aldridge (1982: 217–22), Tunstall (1998: sect. 4), Matthewson (2001: 176–82), Beghelli (1997), and Beghelli & Stowell (1997: 98–104). The latter two papers also contain a comprehensive syntactic proposal that addresses many of the facts above.

Having surveyed the differences between *each* and *every*, I will set them aside and focus on their common core. The analysis I will adopt is closely related to those in Kratzer (2000), Ferreira (2005), and Thomas (2015), but it is more concise than these analyses and makes the connection to the D operator prominent:

- (125) $\llbracket \text{each} \rrbracket_{\text{determiner}} = \llbracket \text{every} \rrbracket$
 $= \lambda P \lambda \theta \lambda V \lambda e. \theta(e) = \bigoplus P \wedge \llbracket D_\theta \rrbracket(V)(e)$
 $= \lambda P \lambda \theta \lambda V \lambda e. \theta(e) = \bigoplus P \wedge e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$

I assume that the determiner combines first with a nominal (of type $\langle e, t \rangle$) and then with a θ -head. Unlike its adnominal and adverbial counterparts, determiner *each* is not coindexed with anything because it is not a distance-distributive item. The thematic relation is contributed by the θ -head. The result has the modifier type $\langle \nu t, \nu t \rangle$ and is ready to combine with the verb phrase or other verbal projection V . A sample noun phrase denotation is shown in (126), and a sample sentence in (127).

- (126) $\llbracket [\text{agent}] \text{ every boy} \rrbracket$
 $= \lambda V \lambda e. * \text{agent}(e) = \bigoplus \text{boy} \wedge e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$

(Takes an event predicate V and returns a predicate that holds of any event e whose agent is all the boys and which consists entirely of events that are in V and whose agents are individual boys.)

- (127) $\llbracket [\text{agent}] \text{ every boy carried three suitcases} \rrbracket$
 $= \exists e. * \text{agent}(e) = \bigoplus \text{ boy } \wedge$
 $e \in * \lambda e' \left(\begin{array}{l} * \text{carry}(e') \wedge |* \text{theme}(e')| = 3 \wedge * \text{suitcase}(* \text{theme}(e')) \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$

This says that there is an event e whose agent is all the boys and which consists entirely of carrying events whose agents are individual boys and whose themes are sums of three suitcases. From this and from the assumption that the agent role is a sum homomorphism, we can conclude that every boy carried three suitcases.

Treating *every* and *each* as involving distributivity operators in a Neo-Davidsonian event-semantic framework like the present one avoids problems in connection with examples like the following (Taylor 1985, Schein 1993, Kratzer 2000, Ferreira 2005, Champollion 2010a, Thomas 2015):

- (128) a. Unharmoniously, every organ student sustained a note on the Wurlitzer for sixteen measures. (Schein 1993)
 b. In a complete lack of harmony, each monk started to sing the Kyrie in a different mode. (Thomas 2015)

The event modifiers in these examples need access to the sum of the events whose agents are the individuals quantified over by *every* and *each*. For reasons analogous to those discussed in Section 8.5 in connection with leakage, the modified event must not be larger than that sum (Ferreira 2005: 23). The entry in (125) allows us to analyze (128a) correctly and concisely:

- (129) $\llbracket (128a) \rrbracket = \exists e. \text{unharmonious}(e) \wedge * \text{agent}(e) = \bigoplus \text{organ.student}$
 $\wedge e \in * \lambda e' \left(\begin{array}{l} * \text{sustain}(e') \wedge \text{note}(\text{theme}(e')) \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$
 (There is an unharmonious event e whose agent is all the students and which consists entirely of note-sustaining events whose agents are individual students.)

Building on insights by Schein (1993) and Kratzer (2000), we can also use the sum event to account for cumulative readings of *every* and *each* such as the ones available in (130a) and (130b).

- (130) a. Three video games taught every quarterback two new plays. (Schein 1993)
 b. Three copy editors caught every mistake (in the manuscript). (Kratzer 2000)
 c. Two farmers sold each sheep to one customer. (Thomas & Sudo 2016)

Such configurations cause problems for the traditional analysis in (120), which does not provide us with access to this sum event. Just as the adverbial modifier *unharmoniously* needs access to the sum of all the individual events in (128a), so do the subject noun phrases in (130). For example, the cumulative reading of (130b) can be paraphrased roughly as ‘There is a sum of mistake-catching events, whose agents sum up to three copy editors, and every mistake was caught in at least one of these events’ (Schein 1993, Kratzer 2000, Champollion 2010a). In this reading, the relationship between the two verbal arguments is cumulative and symmetric. There is no entailment that any mistake was caught by more than one copy editor, as would be expected if *every mistake* took scope either above or below *three copy editors*. My analysis of this reading is as follows.

$$(131) \quad \llbracket (130b) \rrbracket = \exists e. |*agent(e)| = 3 \wedge *copy-editor(*agent(e)) \wedge \\ *theme(e) = \bigoplus mistake \wedge e \in *\lambda e' \left(\begin{array}{l} *catch(e') \wedge \\ Atom(theme(e')) \end{array} \right)$$

This formula says that there is an event whose agents sum up to three copy editors, whose themes sum up to all the mistakes, and which consists of catching events with atomic themes. That these themes are individual mistakes follows from cumulativeness of thematic roles.

It appears that *every* can never enter a cumulative relation with an argument in its syntactic scope (Champollion 2010a). For example, (132) does not have a cumulative reading, in contrast to (130b) (Kratzer 2000).

(132) Every copy editor caught 500 mistakes.

Likewise, Bayer (1997) judges (133a) to be “clearly bizarre” because scripts cannot be written more than once, but reports that (133b) has a reading where every screenwriter in Hollywood contributed to the writing of the movie.

- (133) a. Every screenwriter in Hollywood wrote *Gone with the Wind*.
b. *Gone with the Wind* was written by every screenwriter in Hollywood.

Assuming that *Gone with the Wind* denotes a sum entity, we can represent (133b) as a cumulative reading. Similarly, Zweig (2008) reports that (134a) entails that each game was won by both teams at once, but (134b) has a cumulative reading, in which either team won games and every game was won by only one of the teams.

- (134) a. Every game was won by the Fijians and the Peruvians.
b. The Fijians and the Peruvians won every game.

These facts are in line with what we would expect, since *every* distributes over its syntactic scope but makes the sum event available for arguments or adverbs further up the tree. In (133a) and (134a), the syntactic scope of the argument headed by *every*

is the entire verb phrase. The verb phrase includes the other argument, which is then related as a whole to each of the individual screenwriters or games. As a result, the *every*-phrase takes scope over its coargument and a cumulative reading is ruled out.

- (135) $\llbracket (133a) \rrbracket = \exists e. *agent(e) = \bigoplus screenwriter \wedge$
 $e \in * \lambda e' \left(\begin{array}{l} *write(e') \wedge *theme(e') = \llbracket Gone with the Wind \rrbracket \\ \wedge Atom(agent(e')) \end{array} \right)$
- (136) $\llbracket (134a) \rrbracket = \exists e. *theme(e) = \bigoplus game \wedge$
 $e \in * \lambda e' \left(\begin{array}{l} *win(e') \wedge *agent(e') = \llbracket the Fijians and the Peruvians \rrbracket \\ \wedge Atom(theme(e')) \end{array} \right)$

By contrast, in (133b) and (134b), the syntactic scope of the *every*-phrase only includes the verb. For this reason, it does not distribute over the other argument, and a cumulative reading is possible. Distributing over the verb does not amount to anything much in (134b) since *win* is already distributive on its theme.

- (137) $\llbracket (133b) \rrbracket$
 $= \exists e. *theme(e) = \llbracket Gone with the Wind \rrbracket \wedge$
 $*agent(e) = \bigoplus screenwriter \wedge e \in * \lambda e' \left(\begin{array}{l} *write(e') \wedge \\ Atom(agent(e')) \end{array} \right)$
- (138) $\llbracket (134b) \rrbracket$
 $= \exists e. *agent(e) = \llbracket the Fijians and the Peruvians \rrbracket \wedge$
 $*theme(e) = \bigoplus game \wedge e \in * \lambda e' \left(\begin{array}{l} *win(e') \wedge \\ Atom(theme(e')) \end{array} \right)$

Kratzer (2000) claims that the availability of cumulative readings depends on the thematic role of the coargument of the *every*-phrase. According to her, cumulativity is only possible when the coargument plays the agent role. However, (133b), where the role of the coargument is theme, is a counterexample (Champollion 2010a).

The entry for *each/every* in (125) can be refined in various ways. For example, the subformula $\theta(e) = \bigoplus P$ could be replaced by a contextually supplied variable that specifies the domain of quantification (see e.g. Stanley & Szabó 2000, Matthewson 2001, Schwarz 2009). This variable could be made dependent on another universal quantifier or on a temporal modifier:

- (139) a. Every child ate every apple. (Farkas 1997)
 b. John found a flea on his dog every day for a month. (Zucchi & White 2001)

In order to keep the system simple, I will refrain from adding these refinements here. The temporal case, *every day*, was discussed in Section 8.6.

A related issue is how to analyze the expression *every/each time*. Although I have used it to paraphrase occasion readings of distance-distributive items like *jewels*, we cannot reuse the analysis of those readings. While I have argued that occasion readings involve a nonatomic setting of the granularity parameter, in the case of *each time*

this is not a plausible option. It is natural to assume that expressions involving *each* and mass nouns, like **each mud* or **each silverware*, are ruled out by the atomicity requirement of *each*, which I have implemented by setting its granularity parameter to *Atom*. Since all count nouns are compatible with *each*, even those with unclear individuation criteria such as *twig* or *fence*, it follows that they are all atomic (see Rothstein 2010 and Section 2.6.1). The same reasoning applies to the entities in the denotation of *time*, which is not a mass noun either. On the relevant sense, then, *time* is a nontemporal expression which parcels non-countable entities into countable atoms (Landman 2004: chs 10 and 11). This sense is particularly vivid in sentences like (140):

(140) Each time the bell rings, Mary opens the door.

Such sentences have been analyzed as involving a one-to-one “matching function,” in this case from door openings to bell ringings (Rothstein 1995). On this view, (140) means that every bell-ringing event is the “match” of a different door-opening by Mary. Assuming that the matching function is a sum homomorphism, this type of analysis can be implemented within the present system. Here is the core piece of the derivation:

$$(141) \llbracket \text{each time the bell rings} \rrbracket = \\ \lambda V \lambda e. *match(e) = \bigoplus \llbracket \text{time the bell rings} \rrbracket \wedge \\ e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ Atom(*match(e')) \end{array} \right)$$

For a detailed analysis of this construction, see Landman (2004: chs 10 and 11). That analysis can also be applied to sentences like (142):

(142) When the bell rings, Mary opens one door each time.

A related question is why the *when*-clause that licenses *each time* in (142) cannot license adnominal *each* in (143):

(143) **When the bell rings, Mary opens one door each.*

This is a surprising fact. If *door* involves a one-to-one matching function from door openings to “times the bell rings,” one expects (142) to involve it as well. This matching function has been argued to be a θ -role that is introduced by a silent preposition (Rothstein 1995). If so, the dimension parameter of adnominal *each* in (143) should be able to acquire this matching function as its value via θ -indexing. Setting the granularity parameter to *Atom* should also be possible, since the entities in the denotation of *time* are atomic; otherwise, *each time* would be ungrammatical.

Why, then, is (143) ungrammatical? One possible explanation is that θ -indexing of English adnominal *each* is only available for θ -roles that correspond to (silent or overt) prepositions. Landman (2004: ch. 10) argues contra Rothstein (1995) that sentences like (140) do not involve silent prepositions, and that the matching function expressed by these sentences is not a θ -role but a measure function.

We can also explain why (142), in which the word *time* is present, is grammatical. In this case, the word *each* is not adnominal but a determiner. Since determiner *each* is not distance-distributive, it does not need to be θ -indexed with anything. Therefore the absence of a relevant θ -role does not lead to a violation. The sentence can then be analyzed as in Landman (2004).

Now that we have extended the analysis of distance-distributive items to the English determiners *every* and *each*, we are in a position to explain the generalization discussed in Section 9.2 concerning distance-distributive items and determiners across languages.

9.6 Zimmermann's generalization explained

How can we capture the correlation expressed in Zimmermann's generalization (17)? That is to say, why does a distance-distributive item which can also be used as a distributive determiner lack the occasion reading? One possible explanation is syntactic, as proposed in Zimmermann (2002b). Determiners must agree with their complement; adnominal or adverbial *each* also has a complement, a proform that must acquire its agreement features from its antecedent, which is the antecedent of *each*; only overt antecedents have agreement features; so adnominal/adverbial *each* cannot have a covert antecedent; so it cannot refer to a contextually supplied but not overtly mentioned antecedent such as a salient set of occasions.

This explanation is compatible with the present framework, and it makes the right predictions given the assumption that covert antecedents cannot trigger agreement. However, this assumption is problematic. To mention a simple example, German has grammatical gender. The gender of *Tisch* 'table' is masculine. Knowing this, a German speaker can point to a table and say with reference to it:

- (144) Den hab ich gebraucht gekauft.
 This.acc.m have I used bought
 'I have bought this used.' (German)

But the same speaker cannot point to it and say:

- (145) *Die hab ich gebraucht gekauft.
 This.acc.f have I used bought
 Intended: 'I have bought this used.' (German)

As this example shows, a deictic pronoun in German has to agree in grammatical gender with the gender of the noun phrase that would most aptly describe this antecedent, even though this noun phrase has not been mentioned explicitly.

A similar phenomenon was pointed out for English by Tasmowski-De Ryck & Verluyten (1982). English pronouns agree with their antecedents based on syntactic rather than semantic grounds, as is shown by *pluralia tantum* such as *pants* and

scissors which are syntactically plural but semantically singular. Pronouns show syntactic agreement with their antecedents even when these antecedents are not overtly mentioned:

- (146) a. (John wants his pants that are on a chair and he says to Mary:)
 Could you hand them/*it to me, please?
 b. (Same situation but with a shirt:)
 Could you hand *them/it to me, please?

For a recent discussion of these facts, and an explanation in terms of covert syntactic antecedents that are included in the syntactic structure and c-command the entire sentence in question, see Collins & Postal (2012: ch. 4). In the following I will remain neutral on whether the covert syntactic antecedent should be thought of as being included in the syntactic structure or not.

While Zimmermann's explanation of his generalization seems problematic, its difficulties can perhaps be overcome, and it is by itself not incompatible with the present framework. But in the context of the general theory adopted here, a more straightforward explanation suggests itself. The atomicity requirement of English *each* is of the same kind that Link's D operator provides, as discussed in Section 8.4. The motivation for this requirement, discussed in Section 9.5, can be seen as independent evidence of the atomic distributivity hard-coded in the entry (125) via the D operator. In other words, the distance-distributive item inherits the atomicity requirement of the determiner. This explanation is in line with the notion of parameter settings imported from strata theory as described in Section 9.3. That is, in English, adnominal, adverbial and determiner *each* have essentially identical meanings. Determiner *each* is only compatible with count domains because its granularity parameter is hardwired to the value *Atom*. Adnominal *each* is formally identical to determiner *each*, so it inherits this property.

The German distributive determiner that corresponds to *each* and *every* is *jeder*. The distance-distributive item *jeweils* cannot be used in this position. This is illustrated in sentence (7c), repeated here:

- (147) *Jeder/*Jeweils* Junge hat drei Koffer getragen.
 DIST.sg.m/DIST boy has three suitcases carried
 ‘Every boy has carried three suitcases.’ (German)

This determiner can in turn also be used as an adverbial distance-distributive item. Like English *each*, and unlike German *jeweils*, it can only distribute over individuals, but not over salient occasions.

- (148) Die Jungen haben *jeder* zweimal geniest.
The boys have DIST.sg.m twice sneezed
Available: 'The boys have each twice sneezed.'
Unavailable: 'The boys have sneezed twice on each occasion.'

- (149) *Hans hat *jeder* zweimal geniest.
 Hans has DIST.sg.m twice sneezed
Intended: 'Hans has sneezed twice on each occasion.'

As we see here, the distance-distributive item *jeder* can also be used as a distributive determiner, and it lacks the occasion reading. The distance-distributive item *jeweils* cannot be used as a distributive determiner, and as we have seen before, it has the occasion reading. All this is in line with Zimmermann's generalization. We can account for it by assuming that *jeder*, like *each*, corresponds to the D operator (its granularity parameter can only be *Atom*), while *jeweils* corresponds to the Part operator (its granularity can be set to a nonatomic predicate if it is contextually salient). Concretely, I propose the following denotations for adverbial and determiner *jeder*. They are identical with adverbial and determiner *each* respectively. As for *jeweils*, we have already seen its entry in (81).

$$(150) \llbracket \text{jeder}_\theta \rrbracket_{\text{adverbial}} = \llbracket \text{each}_\theta \rrbracket_{\text{adverbial}} = \llbracket D_\theta \rrbracket = (78)$$

$$(151) \llbracket \text{jeder} \rrbracket_{\text{determiner}} = \llbracket \text{every} \rrbracket \\ = \lambda P \lambda \theta \lambda V \lambda e. \theta(e) = \bigoplus P \wedge e \in {}^* \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right) = (125)$$

The derivation of (147) is exactly analogous to (127). Let me show a derivation of (148). For convenience, and to avoid getting into the difficult question of how to count events, I represent *zweimal* 'twice' as an unanalyzed intersective predicate of sum events. A more sophisticated analysis is found in Landman (2004: chs 10 and 11).

$$(152) \llbracket [\text{agent}] \text{ Die Jungen} \rrbracket = \lambda e. {}^* \text{agent}(e) = \bigoplus \text{boy}$$

$$(153) \llbracket \text{zweimal geniest} \rrbracket = \lambda e. \text{twice}(e) \wedge {}^* \text{sneeze}(e)$$

$$(154) \llbracket \text{jeder}_{\text{agent}} \text{ zweimal geniest} \rrbracket = (151)((153)) \\ = \lambda e. e \in {}^* \lambda e' \left(\begin{array}{c} \text{twice}(e') \wedge {}^* \text{sneeze}(e') \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$$

$$(155) \llbracket (148) \rrbracket = \exists e. e \in (152) \cap (154) \\ = \exists e. {}^* \text{agent}(e) = \bigoplus \text{boy} \wedge e \in {}^* \lambda e' \left(\begin{array}{c} \text{twice}(e') \wedge {}^* \text{sneeze}(e') \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$$

This says that there is a sum event whose agents sum up to the boys, and which consists of sneezing-twice events with atomic agents.

As a *Semantics and Pragmatics* reviewer notes, *jeweils* and *jeder* are morphologically related. They are both built around the distributive item *je*, which also functions as an adnominal distance-distributive item (Link 1998b, Zimmermann 2002b). On the present account, the underlying semantics of *jeder* and *jeweils* is related via the common core of the D and Part operators, discussed in Section 8.4. An interesting question is whether we can explain that the two of them denote related but different

distributivity operators. A starting point might be the observation that the morpheme *weil* in *jewels* is related to the noun *Weile* ‘timespan, while.’ However, the reviewer notes that the morpheme *je* is also found in words with quite distinct meanings, such as *nie* ‘never,’ *jeglich* ‘any kind,’ and *je . . . desto* ‘the . . . the’ (as in *the bigger the better*). As we can see, a common morphological core does not necessarily imply identical meanings. On these questions see also Zimmermann (2002b), who argues that *weil* is a proform; in terms of the present account, it might be the part of *jewels* that is anaphoric on the variable *C*.

Having seen how Zimmermann’s account of his generalization differs from the present one, it is time for a broader comparison of the two systems.

9.7 Previous work: Zimmermann (2002b)

The most detailed semantic account of *jewels* and *each* is offered in Zimmermann (2002b). I summarize and review it here. Other descriptions and critical discussions are found in Blaheta (2003), Dotlačil (2012), and Przepiórkowski (2015). My criticism of Zimmermann’s integrated syntactic and semantic account is limited to its semantic component. I do not take issue with its syntactic component.

Zimmermann takes adnominal *each* and *jewels* to be prepositional phrases that are only partially pronounced, but this aspect does not really influence the semantic composition. The meaning of *each*, or more precisely of the prepositional phrase that is supposed to contain it, is as follows (Zimmermann 2002b: 210). While the relevant discussion is actually about adnominal *jewels*, it carries over to adnominal *each* without changes, so I present it in terms of *each* for clarity.

$$(156) \quad \llbracket \text{each}_{i,j} \rrbracket \quad \quad \quad (\text{Zimmermann}) \\ = \lambda P. \forall z [z \in Z_i \rightarrow \exists x [P(x) \wedge {}^*R_j(z, x)]]$$

This meaning is a property of predicates that holds of a given predicate *P* iff every member of a certain plurality *Z_i* stands in the pointwise algebraic closure **R_j* (see Section 2.3.4) of a certain relation *R_j* to some entity of which *P* holds. In this definition, *Z_i* and *R_j* are free variables that are assumed to be coindexed, respectively, with the antecedent of *each* and with the relation that holds between the host phrase of *each* and its antecedent. That relation is typically denoted by the verb. Take sentence (14), repeated here as (157) with the coindexation added. Here *P* is the denotation of *two monkeys*, *Z* is coindexed with *the boys*, and *R* is coindexed with *saw*.

$$(157) \quad [\text{The boys}]_i \text{ saw}_j \text{ two monkeys each}_{i,j}.$$

Here is how this sentence would be analyzed by Zimmermann (2002b). First, the entry for *each* is applied to *two monkeys*, which is taken to denote a predicate of sum individuals that I will represent here by the shorthand *two-monkeys*. This results in an open proposition with two free variables:

$$(158) \quad \llbracket \text{two monkeys each}_{i,j} \rrbracket \\ = \forall z[z \in Z_i \rightarrow \exists x[\text{two-monkeys}(x) \wedge *R_j(z, x)]]$$

The next steps involve λ -abstracting over the free variables, via a rule that Zimmermann calls “index-triggered λ -abstraction,” a variant of a rule which has been proposed for configurations when a type mismatch makes function application impossible (Bittner 1994: 69).

(159) **Index-triggered λ -abstraction** (Zimmermann 2002b: 217)

If the semantic types of a proposition-denoting expression α and its syntactic sister β do not match, and if $\llbracket \alpha \rrbracket$ contains a free variable u_i that shares an index ‘i’ with β , λ -abstraction in $\llbracket \alpha \rrbracket$ over index ‘i’ is licensed, and $\lambda u_i. \llbracket \alpha \rrbracket$ is a value for α .

This rule allows a constituent with a free variable in it to combine with another constituent that is coindexed with that variable. For example, in (157), the constituent *two monkeys each*_{i,j} has the free variable j in it, which carries the same index as the constituent *saw*_j.⁹⁷ Since the two constituents are sisters, index-triggered λ -abstraction applies, with the result as shown in (160), as discussed in Zimmermann (2002b: 226).

$$(160) \quad \lambda R_j. \forall z[z \in Z_i \rightarrow \exists x[\text{two-monkeys}(x) \wedge *R_j(z, x)]]$$

Zimmermann takes the classical Davidsonian view on verb meaning, under which n -ary verbs denote $n+1$ -ary relations between arguments and events (see Section 2.7). He tentatively proposes that the event argument can “at least sometimes” (p. 226) be saturated inside the verb phrase by existential closure. This means that the verb *saw* can have the right type to combine with (160), as shown below in the derivation taken from Zimmermann (p. 227):

$$(161) \quad \llbracket \text{saw}_j \rrbracket = \lambda y \lambda x. \exists e[\text{see}(x, y, e)]$$

$$(162) \quad \llbracket (160) \rrbracket (\llbracket (161) \rrbracket) \\ = \forall z[z \in Z_i \rightarrow \exists x[\text{two-monkeys}(x) \wedge \exists e[*\text{see}(z, x, e)]]]$$

The result of the computation, in (162), is another open proposition. The last step in the derivation is to combine this with the antecedent, *the boys*_i, in another instance of index-triggered λ -abstraction. The result is as follows:

$$(163) \quad \forall z[(z \in \bigoplus \text{boy}) \rightarrow \exists x[\text{two-monkeys}(x) \wedge \exists e[*\text{see}(z, x, e)]]]$$

This formula says that for every boy there exists a sum of two monkeys such that the boy saw the monkeys. This is an accurate rendering of the truth conditions of sentence (157).

⁹⁷ The need to identify free variables inside constituent denotations poses a challenge for compositionality. To overcome it, Zimmermann (2002a: 336) suggests using partial assignment functions.

In Zimmermann's system, the denotation of the host phrase of adnominal *each*, given in (158), is of type t . This means that the only way it can combine with other constituents is via index-triggered λ -abstraction. The only two indices that can trigger this operation are the ones on Z and R . The values for these two variables will therefore always be provided by the two constituents which are closest to the host phrase. Put another way, Zimmermann's system requires the host phrase of adnominal *each* to be adjacent either to its antecedent or to the constituent that denotes the relation between the two. Whatever intervenes between the host phrase of *each* and its antecedent will give its value to R .

Zimmermann (2002b: 240f.) justifies this adjacency requirement by noting that *jewels* cannot distribute over individual-denoting noun phrases in a higher clause:

- (164) *Die Verkäufer_{*i*} sagen, dass Peter jeweils_{*i*} einen Ballon gekauft hat.
 the store.clerks say, that Peter DIST a balloon bought has
 Intended: 'Each store clerk said that Peter had bought a balloon.'

(Literally: *'The store clerks said that Peter had bought a balloon each.')

(Zimmermann 2002b: 241)

As discussed earlier, this clausemate condition can also be explained by other means. As long as this locality constraint is sufficiently stringent, there is no need for an additional semantic adjacency requirement on top of it. Regardless of how the clausemate condition is implemented, Zimmermann's adjacency requirement is not only redundant in those cases where it agrees with it, it is also too strong where it goes further. This is arguably the case in (165) (Blaheta 2003: 42):

- (165) Alex and Sasha lifted a piano with two jacks each.

In (165), it is possible that only one piano was lifted in a collective event. As Blaheta puts it, the phrase *with two jacks each* "needs to distribute itself in some fashion over each member of the subject, *without* making the verb phrase itself distribute!" Since the host phrase of *each* is not adjacent to its antecedent, it is not obvious how to analyze this configuration in Zimmermann's system. Blaheta leaves (165) as an open problem for his own account as well, which is closely related to Zimmermann's, and he conjectures that event semantics may hold the key to the solution.

This conjecture is correct. The compositional derivation of (165) is similar to the one in Section (86), except that *each* is θ -indexed with the role of the subject, skipping the direct object:

$$(166) \llbracket \text{each}_{\text{agent}} \rrbracket = \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$$

$$(167) \llbracket \text{with two jacks} \rrbracket = \lambda e. \text{two-jacks}(*\text{instrument}(e))$$

- (168) $\llbracket \text{with two jacks each}_{\text{agent}} \rrbracket$
 $= \lambda e. e \in * \lambda e' \left(\begin{array}{c} \text{two-jacks}(*\text{instrument}(e')) \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$
- (169) $\llbracket [\text{theme}] \text{ lifted a piano} \rrbracket$
 $= \lambda e. * \text{lift}(e) \wedge \text{piano}(\text{theme}(e))$
- (170) $\llbracket [\text{agent}] \text{ Alex and Sasha} \rrbracket = \lambda e. * \text{agent}(e) = \text{alex} \oplus \text{sasha}$
- (171) $\llbracket (165) \rrbracket = \exists e. e \in (170) \cap (169) \cap (168)$
 $= \exists e. * \text{agent}(e) = \text{alex} \oplus \text{sasha} \wedge * \text{lift}(e) \wedge \text{piano}(\text{theme}(e))$
 $\wedge e \in * \lambda e' \left(\begin{array}{c} \text{two-jacks}(*\text{instrument}(e')) \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$

This formula entails that Alex and Sasha together lifted a piano, and that each of them was the agent of a part of the lifting event which had two jacks as its instrument. It does not entail that the parts of the lifting events need to be lifting events themselves. This is as it should be, because *lift* is not distributive on its agent position.

Schwarzschild (2014) points out a potential problem for the line of analysis developed here. Suppose that a group of artists build a wall of books on the sidewalk, with each artist putting one book down next to or on top of other books until a wall is built. In this scenario, each artist did something to one book. If we assume that all these events sum up to a building event, sentence (172) should be acceptable and true.

- (172) #The artists built one book each.

The deviant status of (172) can be explained by assuming that the building event is not in fact the sum of the individual events in which artists put down books. For discussion of an analogous problem involving the collective planting of a rosebush, see Kratzer (2007), Williams (2009), and Section 2.5.1. An alternative line of analysis would be to assume that in some cases including (172), the scope of adnominal *each* includes the verb phrase after all. This would raise the question how to delineate the cases in which adnominal *each* does and does not take scope over the verb phrase. I have not adopted this analysis because I do not see an easy way to answer this question.

9.8 Summary

I have suggested the following requirements for a theory of distributivity. First, the synonymy of the adverbial, adnominal, and determiner uses of *each* in English should be captured, ideally by essentially identical lexical entries. Second, the fact that distance-distributive items across languages share some part of their meanings (namely their individual-distributive readings) should be represented, as well as the fact that some of them can also have occasion readings in suitable contexts. Third, the

analysis should clarify the connections between distance-distributive items and distributivity theory more generally, and it should capture the semantic variation across distance-distributive items. Finally, an explanation should be readily available for the crosslinguistic observation that distance-distributive items that can also be used as determiners can only distribute over individuals (Zimmermann's generalization).

I have addressed these issues in the following way. Distance-distributive items across languages are in essence overt versions of Link's D and Schwarzschild's Part operators. The synonymy of the determiner, adnominal, and adverbial uses of *each* in English is captured by the fact that they are all derived from the D operator. I have represented the fact that distance-distributive items across languages share some part of their meanings by deriving them from related distributivity operators (Link's or Schwarzschild's), which differ from each other in their parameter settings and whether they require a formal link to their antecedents. On the theory presented here, distance-distributive items display the same parametric variation as covert distributivity operators do, not only insofar as nonatomic distributivity is concerned, but also insofar as the ability is concerned to target different thematic roles or time. While the syntactic variation among distance-distributive items is due to constraints on formal links that the present theory does not aim to capture, the semantic variation is captured by restrictions on parameter settings. One type of element, exemplified by English *each*, is hard-wired for distribution over atoms; the other one, exemplified by German *jeweils*, also allows distribution over nonatomic contexts. Zimmermann's generalization is explained by the natural assumption that distance-distributive items are formally identical to distributive determiners and therefore inherit their inability to distribute over nonatomic domains, no matter if these domains are mass or temporal.

There are multiple ways in which one could extend the present framework. Cao Yu (p.c.) brought my attention to sentences with multiple distributors such as *the one-on-one tutoring costs \$100 per person per hour*. The semantics of *per person* could be given in essentially the same terms as *one person each*. One relevant advantage of the present framework in this connection is that the distributivity operators introduced in Section 8.4 are stackable: each of them has the same input and output type, namely $\langle v, t \rangle$. This is an advantage compared to Link's and Schwarzschild's original formulation of their distributivity operators, which is not stackable because it returns a truth value. Yenan Sun and Ziren Zhou (p.c.) suggest that the present account may be applicable to the Chinese nonatomic distributor *dou* (e.g. Lin 1998) and to the atomic distributor *ge*. Finally, the framework described here could be extended to *each other* and related reciprocals (LaTerza 2014a).

Taken together with its predecessors, this chapter suggests the following general picture of distributivity. No matter whether distributivity is introduced by an overt or by a covert element, it always involves a certain domain that contains the individuals or the material to be distributed over, and a certain size or granularity that specifies how finely the relevant predicates are distributed. When the domain in question is a

count domain, for example when we distribute over people or objects, then it is always possible to distribute over these objects one by one. When the domain in question does not make such atomic units available, as in the case of time or space, two things can happen. Either the element in question does not allow distribution over such nonatomic domains, for example because it is incompatible with noncount domains to begin with, or else it looks for a salient cover or set in the context, such as a salient set of temporal locations. Those distributive items that can do this in principle can also do this in count domains even though atoms are available.

Collectivity and cumulativity

10.1 Introduction

One of the challenges for any theory of distributivity concerns the relationship between the two words *every* and *all*. In certain contexts, they can be substituted for one another with no discernible semantic effect:

- (1) a. All the students smiled.
b. Every student smiled.
- (2) a. #All the students are a group of five.
b. #Every student is a group of five.

In other contexts, they come apart. Only *all* is compatible with any collective predicates in its syntactic scope, and only *all* can license dependent plurals:

- (3) a. All the students met.
b. #Every student met.
- (4) a. All the students read papers. (*at least one per student*)
b. Every student read papers. (*at least two per student*)

While the examples in (1) and (2) seem to motivate a distributive analysis, those in (3) and (4) appear to favor the view that *all* simply introduces a plural referent and does not distribute over it. This would make *all* more similar to a plural definite determiner than to a universal quantifier.

In general, on the distributive view we expect *all* to pattern with *every*, while on the nondistributive view we expect it to pattern with plural *the*. While the oddness of (2a) is left unexplained by the nondistributive view, the other facts seem to support it. Moreover, plural definites can appear with collective predicates and dependent plurals just like *all*:

- (5) a. The students met.
b. The students read papers. (*at least one per student*)

While many theories account for the facts in (1) through (5), none of them account for a startling observation due to Zweig (2009). When a definite plural combines with a

verb phrase that contains a bounded argument such as a numeral, the resulting reading can be distributive or cumulative. With *every* and *all*, however, only a distributive reading is available:

- | | | |
|-----|---|---|
| (6) | a. The students read thirty papers. | ✓ <i>cumulative</i> , ✓ <i>distributive</i> |
| | b. Every student read thirty papers. | * <i>cumulative</i> , ✓ <i>distributive</i> |
| | c. All the students read thirty papers. | * <i>cumulative</i> , ✓ <i>distributive</i> |

In short, it seems that for every fact that suggests treating *all* as nondistributive, there is another one that suggests treating it as distributive.

The goal of this chapter is to resolve this tension and account for all the facts above, including Zweig's observation. My guiding intuition is that *read papers* is to *read thirty papers* what the atelic predicate *eat apples* is to the telic predicate *eat thirty apples*. But this equivalence comes with a twist: While atelicity is stratified reference along some spatiotemporal dimension, distributivity is stratified reference along some thematic role. While *for*-adverbials test for atelicity, *all* tests for distributivity.

In Chapter 4, I have modeled distributive predicates as having stratified reference down to atoms. Here I will argue that certain collective predicates including *meet* distribute down to sum entities which are larger than atoms, while others including *be a group of five* do not distribute down at all. I will argue that this justifies treating *all* as a distributive item, and that stratified reference helps capture the relevant semantic distinction between these two kinds of predicates. I will show that this view explains why phrases that come with stratified-reference requirements, such as those headed by *all*, cannot participate in cumulative readings with bounded noun phrases in their syntactic scope.

This chapter proceeds as follows. Section 10.2 introduces the idea that *all* is a distributive item just like *every* and *each*. Section 10.3 provides evidence coming from those collective predicates which are incompatible with each of these three quantifiers. Those collective predicates that are compatible with *all* are discussed in Section 10.4. Section 10.5 suggests that *all* distinguishes between these two types of predicates via a stratified-reference constraint which only certain collective predicates satisfy. Section 10.6 shows how this constraint accounts for cases in which *all* appears to take away the collective interpretation of a predicate that can in principle be construed collectively as well as distributively. Section 10.7 focuses on the differences between *all* and *every/each*, and shows that they can be modeled by different values of the granularity parameter. Section 10.8 discusses the limited availability of cumulative readings in sentences involving *all* and *for*-adverbials, and uses stratified reference to account for this fact. Section 10.9 discusses some problems that the present account leaves open, and Section 10.10 concludes.⁹⁸

⁹⁸ The core of the proposal in this chapter was previously published in Champollion (2015c). Sections 10.3 and 10.4 build on Champollion (to appear), which provides an overview of the distinction between the two types of collective predicates. Section 10.8 has been expanded from Champollion (2016b).

10.2 All and maximality

The first set of problems addressed in this chapter concerns the semantic behavior of plural prenominal and adverbial *all*. The status of this word is a challenge for any theory of distributivity. As we have seen, *all* behaves similarly to distributive items like *each* and *every* in some respects, but in other respects it behaves decidedly differently from them. A theory of distributivity must be flexible enough to be applicable to *each* as well as to *all*. As is common in the literature, I do not distinguish here between adverbial *all*, prenominal *all*, *all the*, and *all of the*. For a discussion and a theory of some of the differences between the prenominal items that is compatible with the present account, see Matthewson (2001). I use *all the N* rather than *all N* in the examples in this chapter because *all N* is largely restricted to generic-type contexts, as Matthewson discusses.

I will argue that *all* is a distributive item just like *every* and *each*. The differences between them are due to whether they require predicates to be distributive all the way down to atoms or only down to subgroups. The granularity parameter of stratified reference provides the means to specify this difference.

Historically, the first argument for treating *all* as a distributive item came from the difference in interpretation between *all* and definite plurals. This difference has been called the “maximizing effect” of *all* (Dowty 1987) and, conversely, the “nonmaximality” of definite plurals (Brisson 1998, 2003). Consider for example this minimal pair from Link (1983):

- | | | |
|-----|-------------------------------------|-------------------|
| (7) | a. The children built the raft. | <i>nonmaximal</i> |
| | b. All the children built the raft. | <i>maximal</i> |

As Link notes, “in [(7b)] it is claimed that every child took part in the action whereas in [(7a)] it is only said that the children somehow managed to build the raft collectively without presupposing an active role in the action for every single child.” In other words, (7a) tolerates exceptions in a way that (7b) does not. Link proposes to account for this difference in meaning by giving *all* a translation as a universal quantifier which distributes the property of taking part in building the raft over all individual parts of the totality of children. The difference between (7a) and (7b) is accounted for by assuming that the translation of (7a) does not contain this universal quantifier. On this style of analysis, *all the children* but not *the children* involves distributivity. Link himself only discusses *all* as a side topic in this paper, but his idea underlies the influential analysis of Dowty (1987).

To avoid confusion in the following, let me point out right away that I do not consider the argument from maximality compelling. I present it here because it has been historically important, as it led Link (1983) and following him Dowty (1987) to treat *all* as a distributive item. While I believe that *all* is a distributive item, I do not believe that maximality establishes this fact. As Brisson (1998, 2003) observes,

the maximality/nonmaximality opposition does not correlate with the distributive/collective opposition. Sentences (8a) and (8b) both have a distributive reading, on which each raft was built by a boy, and they also both have a collective reading, on which one raft was built in a coordinated action.⁹⁹ Sentence (8a) tolerates exceptions on both readings, but sentence (8b) presupposes that every boy became involved on both the distributive and on the collective reading (to the extent that this reading is available).

- (8) a. The children built a raft. ✓ *distributive nonmaximal*, ✓ *collective nonmaximal*
 b. All the children built a raft. ✓ *distributive maximal*, %*collective maximal*

As this example shows, the “maximizing effect” of *all* is always present, even when the sentence in which it occurs is interpreted collectively. Likewise, the “nonmaximality” of definite plurals is always present, even when the sentence is interpreted distributively. Therefore, maximality is neither a sufficient nor a necessary condition for distributivity, and it seems problematic to me to conclude on this basis alone that *all* is a distributive item.

I will have nothing to say about the maximizing effect of *all the boys* and the nonmaximality of *the boys*; for recent discussions, see Malamud (2012) and Križ (2016). The translations I will give for these noun phrases ignore the effect. In this respect, my theory contrasts with accounts by Dowty (1987) and Brisson (2003), which use maximality as a way to explain other properties of *all*, such as its inability to license collective readings with certain predicates (as discussed in the next section). These accounts face the problem that *all* shares these properties with other quantifiers in which a maximality effect does not obtain at first sight. According to Winter (2001), *all* patterns with *most of the*, *exactly four*, *at least twelve*, *many*, *few*, and other plural strong quantifiers in terms of its incompatibility with certain collective predicates. It is difficult to see how an item like *few* or even *most* could be claimed to involve a maximality effect in any meaningful way.¹⁰⁰ For this reason, I do not use maximality to explain the properties of *all* which are of interest here. Each of the claims I make about *all* in the following seems to extend to *most of the* (see also Nakanishi & Romero 2004). However, other strong quantifiers such as *exactly three*

⁹⁹ Dowty (1987) reports for similar sentences that some speakers find collective readings more natural if the word *together* is added. I come back to this point in Section 10.4. As we will see, from the point of view of the theory developed here, the fact that some speakers can access a collective reading for sentence (8b) is surprising, and its unnaturalness for other speakers is expected. I do not discuss the semantics of collectivizing adverbials such as *together* in this book, and I do not account for their effect on the status of sentences like (8b). See also Section 10.9 for further discussion of open problems in connection with *together*.

¹⁰⁰ Sven Lauer (p.c.) suggests that items like *most*, as opposed to definite plurals, might involve a reduction of pragmatic slack in the sense that Lasnik (1999) suggests for *all*. It is not clear to me how slack reduction can be diagnosed in *most*, or how it could be linked to the other properties of *all* under consideration in this chapter. Lasnik (1999) himself does not provide such a link and refers to the theory of Dowty (1987) to account for these properties.

pattern differently, for example because they can take part in cumulative readings (Brasoveanu 2013).

10.3 Numerous-type predicates

Having discarded maximality effects, let me now introduce what I believe to be a convincing argument that *all* is a distributive item. This argument comes from a class of predicates with respect to which *all* behaves analogously to the uncontroversially distributive items *every* and *each*. This class was introduced in Section 4.2.3 under the name of *numerous-type predicates*. The predicate *be a group of five* in (2) belongs to this class. While these predicates can give rise to collective interpretations together with definite plurals and other noun phrases, they cannot be interpreted collectively when *all* or *every/each* are present.

This category has also been called *purely collective predicates*, *pure cardinality predicates* (Dowty 1987), and *genuine collective predicates* (Hackl 2002). It roughly corresponds to the *atom predicates* of Winter (2001, 2002), but that class is larger: as discussed in Section 4.2.3, Winter also includes run-of-the-mill distributive predicates like *smile*, which do not have collective interpretations to begin with. Following Dowty (1987), I do not include distributive predicates in the class of *numerous-type predicates*. Many of the facts I discuss in this section were first observed by Kroch (1974) based on related observations by Dougherty (1970, 1971), and independently by Dowty (1987); see also Moltmann (1997: 128f.).

In connection with definite plurals, *numerous-type predicates* easily give rise to collective interpretations. Indeed, the collective interpretation is often the only one available. For example, in (9a), the predicate *be numerous* can only be understood as applying collectively to the ants in the colony, because there is no sense in which an individual ant can be numerous. If a distributive interpretation is available at all, it surfaces when the definite plural is headed by a group noun such as *committee*, *army*, or *family* (Winter 2001). In such cases, some speakers find the sentence acceptable, and in that case it exhibits distributive entailments (Kroch 1974: 194).¹⁰¹ For example, (9b) can mean either that each of the armies taken by itself had many members or that the number of armies was large.

- | | | |
|-----|--|---|
| (9) | a. The ants in the colony were numerous. | * <i>distributive</i> , ✓ <i>collective</i> |
| | b. The enemy armies were numerous. | % <i>distributive</i> , ✓ <i>collective</i> |

The distributive item *each* (and its relative *every*, which behaves analogously to it) only allows the distributive interpretation of a predicate of this type. When there is no such interpretation in the first place, the sentence becomes unacceptable altogether.

¹⁰¹ For other speakers, *numerous* cannot be predicated of entities in the denotation of group terms. See Section 4.2.3.

This is the case in (10a) because an ant cannot be numerous, and I expect it to be the case in (10b) for those speakers for whom *numerous* does not apply to group entities.

- (10) a. *Each/every ant in the colony was numerous. **distributive*, **collective*
 b. %Each/every enemy army was numerous. %*distributive*, **collective*

The effect of *all* on this type of predicate is identical to the effect of *each*: if the sentence is acceptable at all, it only has a distributive interpretation. For example, sentence (11a) is unacceptable, and to the extent that sentence (11b) is acceptable, it can only be interpreted as saying that every enemy army had many members:

- (11) a. *All the ants in the colony were numerous. **distributive*, **collective*
 b. %All the enemy armies were numerous. %*distributive*, **collective*

Sentence (11b), if acceptable, is synonymous with sentence (10b), whose only interpretation is distributive.

Other examples of the *numerous*-type class include *be politically homogeneous*, *be a motley crew*, *suffice to defeat the army* (Kroch 1974), *be a large group*, *be few in number*, *be a couple* (Dowty 1987), *be denser in the middle of the forest* (can be said of trees, Barbara Partee p.c. via Dowty 1987), *pass the pay raise*, *elect Bush*, *return a verdict of 'not guilty'*, *decide unanimously to skip class*, *eat up the cake*, *finish building the boat* (Taub 1989), *be too heavy to carry* (Brisson 1998), *be a good team*, *form a pyramid*, *constitute a majority*, *outnumber* (Winter 2001). However, not all of these predicates behave alike, and some contexts may improve their ability to occur in sentences with *all* (Champollion to appear):

- (12) a. Some of the boys were crying, but eventually (and after much discussion), all the boys formed a (nice) pyramid.
 b. There was a lot of discussion, but eventually, all the boys decided unanimously to skip class.
 c. I know it sounds kind of crazy but in fact all the weapons in this little village would suffice to defeat the US Army.

According to a reviewer of Champollion (to appear), these examples were judged acceptable on a collective construal by native speakers of American English. This could either be taken to suggest that these predicates are classified as *gather*-type rather than *numerous*-type in the grammars of those speakers, or that the word *all* is used in a non-standard sense that lacks its usual restriction. Neither of these possibilities can be easily accounted for, and I will set this phenomenon aside in the remainder of this chapter.

The parallel between (10) and (11) motivates treating *all* in analogous terms to the canonical distributive items *each* and *every*. Before presenting such a treatment, let me turn to a class of predicates which seem to provide evidence that *all* is not a distributive item, thereby presenting a challenge to any parallel treatment of *each* and *all*.

10.4 Gather-type predicates

Gather-type predicates can lead to distributive and collective interpretations, even in the presence of the word *all*. They were introduced in Section 4.2.3 along with *numerous-type* predicates. The predicate *meet* in example (3) belongs to this class. As shown in Table 4.1, both *numerous-type* and *gather-type* predicates are subtypes of collective predicates, in the sense that their distributive reading is only available if the subject argument involves reference to group individuals such as committees. The difference between *gather-type* and *numerous-type* predicates concerns their interpretation. We have seen that the collective interpretation of *numerous-type* predicates is blocked both by *every/each* and by *all*, even when it is the only available interpretation. The collective interpretation of *gather-type* predicates is also blocked by *every/each*, but it is not blocked by *all*:

- | | | |
|------|---|------------------------------|
| (13) | a. All the students gathered in the hall. | *distributive, ✓ collective |
| | b. *Each student gathered in the hall. | *distributive, *collective |
| (14) | a. All the committees gathered in the hall. | ✓ distributive, ✓ collective |
| | b. Each committee gathered in the hall. | ✓ distributive, *collective |

The observation that some collective predicates (namely the *gather-type* ones) are compatible with *all* but not with *each* goes back at least to Vendler (1962). The *numerous/gather* opposition has subsequently been discussed in Kroch (1974), Dowty (1987), Taub (1989), Moltmann (1997), Brisson (1998, 2003), Winter (2001, 2002), Hackl (2002), Dobrovie-Sorin (2014), Kuhn (2014), and elsewhere. *Gather-type* predicates have also been called *essentially plural predicates* (Hackl 2002) and *set predicates* (Winter 2001). Other examples of this type of predicate are *be similar*, *fit together* (Vendler 1957), *meet*, *disperse*, *scatter*, *be alike*, *disagree*, *surround the fort*, *summarize* with respect to its object argument (Dowty 1987), *have common interests* (Link 1987b), *form a big group* (Manfred Krifka p.c. via Brisson 2003), *be consistent*, *be compatible*, *hold hands* (Kuhn 2014).

It is not easy to draw the boundary of the class of *gather-type* predicates. If one includes all collective predicates into this class as long as they are compatible with *all*, as in Winter (2001), one ends up with a heterogeneous class, which includes reciprocally interpreted predicates such as *admire each other*, and predicates formed with collectivizing adverbials such as *perform Hamlet together*. Following Dowty (1987) and Brisson (2003), I exclude these predicates from consideration here. Winter furthermore includes any predicate that is compatible both with *all* and with *each* as long as they bring about a difference in truth conditions. This difference cannot always be easily attested. For example, predicates which can be understood distributively as well as collectively, like *build a raft* and *perform Hamlet*, belong to this class so long as their collective interpretation remains available with *all* and can be distinguished truth-conditionally from their distributive interpretation, as in (15). As reported in

case as giving rise to the sub-entailment “has a location which is relatively close, in every direction, to that of another member of the group.”

Taub (1989) suggests that all *gather*-type predicates are activities and accomplishments, while all *numerous*-type predicates are states and achievements. Following her, Brisson (1998, 2003) proposes a syntactic account of the *numerous/gather* opposition that implements this in terms of a silent predicate DO. Brisson assumes this predicate to be present only on activities and on accomplishments (I discuss this proposal further in Section 10.8). However, Taub’s hypothesis seems more like a tendency than a hard-and-fast generalization. For example, the predicate *reach an agreement* is as good an achievement predicate as any other, but it is *gather*-type since it is compatible with *all* on a collective reading:

- (18) All the parties involved reached an agreement.

As discussed by Brisson (2003), predicates like *look alike*, reciprocals, and collectivized predicates constitute counterexamples as well, since they can be used to form *gather*-type states and achievements:

- (19) a. All the dogs look alike.
b. All the boys recognized each other.
c. All the planes arrived together.

Winter (2001) assumes that *gather*-type collective predicates apply to sets, which correspond in his system to mereological sums, while *numerous*-type and distributive predicates apply to atoms, which may be either pure or impure (see Section 2.8). A crucial assumption of his system is that sets and atoms have different types. This makes it difficult for him to treat predicates like *build a raft*. These predicates have distributive as well as *gather*-type collective interpretations, and Winter must assign them two different types. Unlike Winter (2001), I will account for the difference between *numerous*-type and *gather*-type predicates without assuming that there is a type-theoretic distinction between the two.

Let me now turn to my own proposal. Predicates like *gather*, *meet*, and *disperse* are traditionally classified as collective because they do not distribute down to individuals. However, there is a generalized notion of distributivity that applies to all of them: they distribute to subgroups that are small in number. For example, if a group of ten people gathered or dispersed, then smaller subgroups of two or three people within this group also gathered or dispersed. By contrast, other collective predicates, such as *be numerous* or *be a good team*, do not license this kind of entailment, since two people do not count as being numerous and since two people taken from a good team do not typically form a team by themselves (except when the good team consists of only these two people to begin with, as in tennis). Let me generalize this to the following hypothesis. This and the following discussion are based on Kuhn (2014) and

Dobrovie-Sorin (2014), who in turn build in part on an earlier version of this chapter (Champollion 2010b: ch. 9).

(20) **The subgroup distributivity hypothesis**

Gather-type predicates are those that distribute down to subgroups of small cardinality. Informally, if a collective predicate holds of a plural entity *X*, it will be a *gather*-type predicate just in case there is a way to divide *X* into at least two small, possibly overlapping subgroups (usually pairs or triples) such that the predicate applies to each of these subgroups.

Many *numerous*-type predicates are accurately described by this hypothesis. For example, *be numerous* by definition is not true of subgroups that are small in number, and for legal reasons, *return a verdict of 'not guilty'* cannot hold of any subset of any jury of which it holds. Other *numerous*-type predicates that are correctly characterized include *suffice to defeat the army*, *be a group of four*, *pass the pay raise*, *elect Bush*, *decide unanimously to skip class*, *eat up the cake*, *finish building the boat*, *be too heavy to carry*, *be a good team*, *form a pyramid*, *constitute a majority*. However, some others including *be a group*, *be a motley crew*, *be small in number*, or *be politically homogeneous* might be argued to be subgroup distributive and need to be given a separate treatment. I leave this problem open here; a corresponding problem has been noted for *for*-adverbials in connection with “bounded but nonquantized” predicates like *write a sequence of numbers* or *drink a quantity of milk* (Zucchi & White 2001).

The subgroup distributivity hypothesis also correctly characterizes many *gather*-type predicates, including *gather* itself as well as *be similar*, *meet*, *disperse*, *scatter*, *be alike*. In fact, this tendency was already observed by Winter (2001: 223f.).¹⁰² For example, if ten people are similar, then any two of them are similar.

Despite the overall good match of the hypothesis, it should be acknowledged that some predicates that lack subgroup distributivity are nevertheless compatible with *all* to various extents. This concerns the collective readings of predicates like *perform Hamlet*, *carry the piano upstairs*, and *build a raft* on Dowty's dialect, as well as predicates that contain collectivizing adverbials such as *build a raft together* (Brisson 2003) and *build a raft in the same room* (Jeremy Kuhn, p.c.). A possible explanation, based on Brisson (2003), is that such verb phrases include a covert semantically bleached-out *gather*-type event predicate DO and that it is this predicate rather than the entire verb phrase that satisfies the requirements of *all* when present. The dialectal and prosodic differences described earlier might then be attributable to the distribution of this event predicate. I am reluctant to endorse this explanation for reasons explained in Section 10.8, and I leave the problem open here.

¹⁰² Winter does not ascribe subgroup distributivity to *gather*-type predicates in general because he regards the predicate *drink together a whole glass of beer* as a counterexample. Jeremy Kuhn and Todd Snider (p.c.) find the placement of *together* unnatural here. To the extent that this predicate along with similar predicates is acceptable, I must leave it as an open problem here; see also Section 10.8.

In Chapter 4, I captured the distributive/collective opposition at the level of lexical predicates by formulating meaning postulates for distributive predicates. We can see a distributive predicate as one that has stratified reference down to atoms on the appropriate thematic position. Collective predicates do not satisfy stratified reference down to atoms, because none of the atomic parts of these subjects participate in events that satisfy these predicates. We can formulate a modified form of such meaning postulates for those collective predicates that distribute to subgroups. The view of distributivity as stratified reference suggests that not only the dimension parameter but also the granularity parameter should have a role to play. This is indeed the case: it can be used to specify “how distributive” a given predicate is on a given argument position, for example whether it distributes all the way down to atoms or only to small but nonatomic entities. I will formalize meaning postulates with the help of the universal version of stratified reference, which was defined in Section 4.6. In expanded form, this definition looks as follows:

(21) **Definition: Universal stratified reference**

$$\text{SR}_{d,g}(P) \stackrel{\text{def}}{=} \forall x \left[P(x) \rightarrow x \in {}^*\lambda y \left(\frac{P(y) \wedge}{g(d(y))} \right) \right]$$

(*P* has universal stratified reference along dimension *d* with granularity *g* iff any *x* in *P* can be divided into one or more parts in *P* that are each mapped by *d* to something in *g*.)

For present purposes, the property *P* will be the verbal predicate, and the parameter *g* will encapsulate the size of the entities to which the predicate is distributed. Although it might seem natural in the context of *all* and *each* to think of the plural entities over which *x* ranges as groups of people, I will instead let *x* range over the plural events (or eventualities) described by the verbal predicate. Using events throughout makes it easier to draw parallels between the collective/distributive and the telic/atelic opposition. I will use the dimension parameter *d* to mediate between the subevents of *x* and the granularity level *g*. Specifically, *d* will be instantiated with the thematic role of the noun phrase that hosts the determiner in question. This thematic role will map the event in question to the plural entity which is then tested for whether it is sufficiently small.

All *gather*-type predicates can be described by similar meaning postulates. As an example, *meet* satisfies the following meaning postulate (as described in Section 2.5.5, I use $|x|$ for the number of atomic parts of *x*):

(22) **Meaning postulate for *meet***

$$\begin{aligned} &\text{SR}_{\text{agent}, \lambda x. |x|=2}(\lambda e. {}^*\text{meet}(e)) \Leftrightarrow \\ &\forall e \left[{}^*\text{meet}(e) \rightarrow e \in {}^*\lambda e' \left(\frac{{}^*\text{meet}(e') \wedge}{|{}^*\text{agent}(e')|=2} \right) \right] \end{aligned}$$

(Every meeting event can be divided into one or more parts, each of which is a meeting event whose agent is two people.)

Having formalized the distinction between *gather*-type and *numerous*-type predicates, two questions arise: Why do *gather*-type and *numerous*-type predicates behave differently with respect to *all*? And why do *all* and *every/each* behave differently with respect to *gather*-type predicates? I turn to these questions now.

10.5 Why *all* distinguishes between *numerous* and *gather*

I propose that *all* distinguishes between *numerous*-type and *gather*-type predicates because, simply put, it requires verb phrases to apply to smaller parts of anything they apply to (see Kuhn 2014, Dobrovie-Sorin 2014, Champollion 2015c, 2016b). This subgroup distributivity requirement rules out *numerous*-type predicates.

My analysis is based on the idea that *all* imposes a form of the Distributivity Constraint (see Section 4.6) on the verb phrase predicate. The main difference is that the “dimension” involved (the Map, in the terminology introduced in Section 4.4) is not runtime or location as in the case of *for*-adverbials, or some measure function as in the case of pseudopartitives, but the thematic role of the *all*-phrase. In this respect, the analysis of *all* is similar to that of *each* and *every* discussed in Chapter 9. This constraint has the effect that when *all* modifies a verb phrase (either directly as an adverb, or preminally after combining with a bare plural or definite description), any event in the denotation of the verb phrase is required to consist of one or more subevents, or strata, which are in the denotation of the verb phrase and which the thematic role associated with *all* maps to a value that is small in number. This is captured via an appropriate setting of the granularity parameter.

I will formalize this constraint with the help of the universal version of stratified reference. Specifically, if *all* heads a subject noun phrase, I assume that it imposes stratified reference as a presupposition on the verb phrase. This essentially means that *all* makes sure that the verb phrase distributes down at least to sums that are small in number. In some cases including distributive predicates and those which can be understood both distributively and collectively, the verb phrase will in fact distribute down to singular individuals:

(23) All the boys { smiled / arrived / won }.

In other cases, namely the *gather*-type predicates, the verb phrase will distribute down only to sums that are small in number, but not to singular individuals:

(24) All the boys { gathered / met / dispersed }.

Both cases satisfy the requirement of *all*.

Formally, I will leave open what counts as small in number, and I will simply assume that this is determined by a placeholder predicate ε . This predicate takes a comparison class K (either a set or a sum of entities) and an entity x , and returns true just in case x is small (typically small in number) compared to K . For example, $\varepsilon(\bigoplus \text{boy})(x)$ holds

just in case x is small in number compared to the sum of all boys. I assume that ε is not reflexive (nothing counts as small by comparison with itself). This rules out sentences with singular count nouns like **All the boy smiled*, since these nouns are atomic (see Section 2.6.1). I also assume that ε is downward monotonic on its second argument in the sense of having divisive reference (if x is small compared to K , then any part of x is also small compared to K). I assume that when *all* heads a noun phrase that denotes the agent of an event e , its comparison class argument K is instantiated with the plural entity that is the agent of e , and similarly for other thematic roles. Unlike the C parameter of the Part operator discussed in Section 8.3, I do not assume that the ε parameter of *all* is anaphoric; the granularity level of *all* does not need to be salient in context.

One reason I do not commit to a concrete numerical threshold is that *all* can also combine with mass terms, in which case cardinality is not an applicable notion. For example, Winter (2001: 211) observes that *all the sugar* is compatible with a collective predicate that can be classified as *gather*-type in the count domain:

- (25) All the sugar is concentrated in this container.
- (26) a. All the pieces of sugar are concentrated in this container.
b. ?Every piece of sugar is concentrated in this container.

Dobrovie-Sorin (2014) argues based on observations by Bunt (1979), Lønning (1987), and Higginbotham (1994) that *all* imposes an analogous constraint in the mass domain as it does in the count domain. Distributive predicates like *frozen* are compatible with *all*, while those that lack subgroup distributivity are not:

- (27) All the water { is frozen / *weighs one ton }.

That the relevant notion is subgroup distributivity and not distributivity per se was already observed by Carlson (1981), who notes that *All the ground was speckled with leaves* “does not mean that there were absolutely no bare spots (that is even excluded by the meaning of *speckled*), only that there were no bare spots big enough to break a pattern of speckles.” In other words, the minimal-parts problem arises with *all* just like it does with other distributive items.

To account for the behavior of *all*, I assume that it imposes a stratified-reference presupposition after it combines with its thematic role, with a sum-denoting constituent such as *the boys*, and then with a verbal projection:¹⁰³

- (28) $\llbracket \text{all} \rrbracket = \lambda \theta \lambda y \lambda V \lambda e : \text{SR}_{\theta, \varepsilon(\theta(e))}(V).[V(e) \wedge * \theta(e) = y]$

¹⁰³ Following Matthewson (2001), I assume that prenominal *all* combines with an entity of type e , which may be provided either by a definite plural such as *the boys* (possibly mediated by a vacuous preposition of) or by a bare plural such as *boys*, which is interpreted as a kind-referring term (see Section 2.4.2). The entry in (28) only covers prenominal *all*; for adverbial *all*, an analogous entry can be obtained by changing the order of its arguments to match those of the entry for adverbial *each* in Section 4.7.

For *All the boys smiled*, this entry yields the following meaning:

- (29) $\exists e : \text{SR}_{\text{agent}, \varepsilon}(*\text{agent}(e))(*\text{smile}).[*\text{smile}(e) \wedge *\text{agent}(e) = \bigoplus \text{boy}]$
 (There is a smiling event e whose agents sum up to the boys, with the presupposition that *smile* has stratified reference along the *agent* dimension with granularity $\varepsilon(*\text{agent}(e))$.)

The presupposition of (29) expands as follows, where e is any sum event that verifies (29):

- (30) $\text{SR}_{\text{agent}, \varepsilon}(*\text{agent}(e))(*\text{smile})$
 $= \forall e'. * \text{smile}(e') \rightarrow e' \in *\lambda e'' \left(\begin{array}{c} * \text{smile}(e'') \wedge \\ \varepsilon(*\text{agent}(e))(*\text{agent}(e'')) \end{array} \right)$
 (*Smile* has stratified reference along the *agent* dimension with granularity $\varepsilon(*\text{agent}(e))$ (“small in number compared to the agent of e ”) iff any smiling event can be divided into one or more smiling events whose agents are each small in number compared to the agent of e .)

This presupposition is satisfied because *smile* is distributive along its agent dimension (see Chapter 4). That is, the only kinds of smiling events that have sum individuals as their agents are those that consist of several smiling events which each have an atomic agent. In general, assuming that *all the boys* refers to more than one boy to begin with, any predicate that is distributive along the relevant dimension will satisfy the presupposition in (30). To make the presentation simpler, from now on I will assume without loss of generality that *all* is always in agent position, that it modifies the definite plural *the boys*, and that the sentence is interpreted in a model where there are several boys.

For the *gather*-type collective predicate *meet*, the entry in (28) yields a presupposition that is satisfied:

- (31) all the boys met
Satisfied presupposition: $\text{SR}_{\text{agent}, \varepsilon}(*\text{agent}(e))(*\text{meet})$
 $= \forall e'. * \text{meet}(e') \rightarrow e' \in *\lambda e'' \left(\begin{array}{c} \text{meet}(e'') \wedge \\ \varepsilon(*\text{agent}(e))(*\text{agent}(e'')) \end{array} \right)$
 (Any meeting event can be divided into one or more meeting events whose agents are each small in number compared to the agent of e .)

Because *meet* distributes down to subgroups as per the meaning postulate in (22), this presupposition is satisfied even though *meet* does not distribute down to atoms. In general, distributive predicates like *smile*, certain predicates like *win* that are ambiguous between distributive and collective interpretations, and *gather*-type collective predicates will satisfy the presupposition of *all* (Kuhn 2014).

However, those collective predicates that do not have stratified reference down to a small level of granularity will not satisfy the presupposition of *all*. I have suggested

that these are just the *numerous*-type predicates (such as *be numerous*, *be a group of five*, *suffice to defeat the army*). For example, a plurality of soldiers may suffice to defeat the army while small subgroups of this plurality would not suffice. And although the plurality of ants in my kitchen is numerous, small subgroups are not:

- (32) *all the ants in my kitchen are numerous

Failed presupposition: $SR_{agent, \varepsilon(agent(e))}(\text{be numerous})$

$$= \forall e'. \text{numerous}(e') \rightarrow e' \in * \lambda e'' \left(\begin{array}{c} \text{numerous}(e'') \wedge \\ \varepsilon(agent(e))(agent(e'')) \end{array} \right)$$

(Any eventuality of being numerous can be divided into one or more eventualities of being numerous whose agents are each small in number compared to the agent of e .)

Due to the nature of stratified reference, the constraint that *all* imposes does not require the predicate in question to apply to all parts whose agent is small in number, only that there be some way of dividing the whole plural entity into such parts. As Kuhn (2014) observes, some collective predicates that exhibit precisely this kind of subgroup distributivity are indeed compatible with *all*, as expected:

- (33) a. All the pieces of the puzzle fit together.
b. All the boys in the circle held hands.

In typical scenarios that make these sentences true, the predicates *fit together* and *hold hands* do not apply to all pairs of puzzle pieces or to all pairs of boys, but only to pairs of adjacent ones. Even so, stratified reference holds, and the constraint imposed by *all* is satisfied. For example, the constraint in (33b) requires that any hand-holding event can be divided into one or more hand-holding events whose agents are each small in number, and this is indeed the case. In Chapters 5 through 7, this feature of strata theory made sure that examples like *swim laps for an hour*, *push carts all the way to the store for fifty minutes*, and *five feet of snow* are not ruled out. In all these cases, universal quantification over all parts of the event or entity in question would lead to the wrong prediction even if very small parts were excluded.

While the subgroups to which the predicate applies are adjacent in the sentences in (33), stratified reference imposes no adjacency requirement. Therefore, to the extent that a small number of appropriately positioned soldiers can count as surrounding a town, sentence (34) is correctly predicted to be acceptable:

- (34) All of the soldiers in the battalion surrounded the town. (Kroch 1974: 187)

Likewise, example (35) is predicted acceptable because there will generally be a way to subdivide the total amount of salt into parts each of which forms a square, since a square can always be divided into smaller squares:

- (35) All of the salt formed a square on the floor.

(adapted from Dobrovie-Sorin 2014)

To summarize this section, we may say that *all* is an “almost distributive” determiner: it requires distributivity down to a small granularity level, but not all the way down to atoms (which are not even available to begin with in the mass domain). *Gather*-type predicates are “a bit distributive” (they have subgroup distributivity); *numerous*-type predicates are either not distributive at all (*be a group of ten*) or not enough to satisfy the presupposition of *all* (*be numerous*).

A final remark on the notion that what sets *gather*-type predicates apart within the class of collective predicates is that they distribute down to elements of low cardinality. This basic idea is taken in essence from Kuhn (2014) and represents a slight departure from Champollion (2010b), where I assumed instead that collective predicates technically never distribute and that *gather*-type—but not *numerous*-type—collective predicates have groups in the sense of Landman (1989, 1996) as agents. As discussed in Section 2.3.4, groups are mereologically atomic entities that formally model plural individuals and that give rise to certain kinds of entailments about themselves involving collective action or collective responsibility. The atomic nature of groups made sure that the stratified-reference requirement was always vacuously satisfied. One problem with that view was that it did not capture the inference from *A, B, and C met* to *A and B met*. Another problem was that Landman and myself only gave a partial list of what constitutes collective entailments. What the present account has in common with Champollion (2010b), however, is the idea that *all* imposes stratified reference just like *each* (though with different granularity parameter settings), and that both can therefore be seen as distributive items.

10.6 Meaning shift effects triggered by *all*

The account so far predicts that *numerous*-type predicates are never compatible with the presupposition of *all*, and that they should never cooccur. We have already seen in Section 10.3 that this prediction is too strong. For some speakers, the predicate *be numerous* is compatible with *all* when the subject of the sentence is headed by a group noun. This is illustrated by example (11b), repeated here as (36), which is described as having a distributive reading in Kroch (1974: 194):

- (36) All the enemy armies were numerous. ✓ *distributive*, **collective*

This sentence poses two challenges. First, we need to explain why it does not conform to the prediction that *all* and *be numerous* cannot cooccur. Second, we need to explain why it has a distributive reading.

A similar problem is posed by sentences in which the addition of *all* to a predicate results in meaning shift rather than unacceptability. When *all* combines with a predicate that can otherwise be interpreted both distributively and collectively, it sometimes appears to take away the collective reading. Dowty (1987) attributes the following minimal pair to Bill Ladusaw:

- (37) a. The students voted to accept the proposal. ✓ *distributive*, ✓ *collective*
 b. All the students voted to accept the proposal. ✓ *distributive*, **collective*

In order for (37a) to be true, all that is required is that the students as a whole accepted the proposal by voting on it according to whatever are the appropriate rules and procedures. This is a collective reading. Given our background assumptions about such procedures, a simple majority vote is likely to suffice to make the sentence true. In the case of (37b), though, the vote must have been unanimous. This is a distributive reading.

Similarly, Taub (1989) notes that in the presence of *all*, the predicates *win* and *be heavy*, which usually give rise to collective as well as distributive interpretations, can only be read distributively:

- (38) a. The boys won. ✓ *distributive*, ✓ *collective*
 b. All the boys won. ✓ *distributive*, **collective*
- (39) a. The stones are heavy. ✓ *distributive*, ✓ *collective*
 b. All the stones are heavy. ✓ *distributive*, **collective*

We can account for all these cases by assuming that a covert distributivity operator is inserted. This idea is similar to Brisson (2003); unlike that theory, however, I do not assume that the word *all* is licensed by the presence of a distributivity operator. Instead, I build on Section 8.6 and view the distributivity operator as a repair strategy that changes the meaning of a predicate so that it satisfies the presupposition of a distributive item. I assume that the LF for sentences like (36) and (37b) contains a distributivity operator, which is inserted as a means to meet the requirements of *all* when the predicate would not otherwise meet these requirements. Because the operator is defined in terms of stratified reference, its output meets the requirements imposed by *all* (see the Appendix for a proof). The operator shifts the predicate to a distributive interpretation; this accounts for the meaning shift.

My formulation of the distributivity operator is repeated here from Section 8.2:

(40) **Definition: Event-based D operator**

$$\llbracket D_\theta \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$$

The granularity parameter of this operator is hardwired to the value *Atom*. As for its dimension parameter, it can be instantiated with any thematic role. I will assume that it is instantiated with *agent*. This may not be the best choice in the case of *numerous*, but since the identity of the thematic role does not matter, I will stick to it for convenience (see Section 2.5.1).

When this operator is applied to *be numerous*, it returns a predicate that holds of any event *e* which consists of one or more events that are in *be numerous* and whose agent

is an atom. This predicate applies to fewer events than *be numerous* does. For example, imagine a model in which there are at least the following entities: one thousand boys and the Golden Horde (a famous army). The boys have nothing to do with the Golden Horde. The members of the Golden Horde might also be in the model, but they are not of interest here. The word *army* is a group noun, and the Golden Horde is one of its referents, so the Horde is represented as an atom, call it *g* (see Section 2.6.4). Assume that the sum of all boys in the model qualifies as numerous, and the Golden Horde does too. Then the predicate *be numerous* applies at least to two events e_1 and e_2 such that the agent of e_1 is \oplus boy and the agent of e_2 is *g*. However, the predicate $D_{agent}(\llbracket \text{be numerous} \rrbracket)$ only applies to e_2 and not to e_1 , because only e_2 consists of one or more events (specifically, of one event) whose agent is an atom.

In this way, the distributivity operator acts as a filter on the predicate *be numerous*: its output only contains those events that have as their agents the referents of group nouns such as armies, committees, and so on, or which are built up from such events. In other words, every event in the denotation of the predicate $D_{agent}(\llbracket \text{be numerous} \rrbracket)$ can be divided into one or more parts each of which is in the denotation of $D_{agent}(\llbracket \text{be numerous} \rrbracket)$ and has an atomic agent. Therefore, $D_{agent}(\llbracket \text{be numerous} \rrbracket)$ satisfies the presupposition of *all* even though *be numerous* does not. This explains why sentences like (36), to the extent that they are acceptable, only have a distributive interpretation: they contain the covert operator D_{agent} . I come back to this filtering property of the D operator in Section 10.8.

The explanation for the contrast between (37a) and (37b) is analogous: The predicate *vote to accept the proposal* applies to individuals (in the sense of casting a ballot) as well as to groups (in the sense of ratifying a proposal). In (37a), by default the predicate is applied directly to the subject, and the collective reading emerges (see Section 2.8). In (37b), this derivation is not available because the predicate does not satisfy the stratified-reference requirement of *all*: if a group accepts a proposal by a simple majority vote and a sizeable subgroup votes against it, the predicate will fail to distribute within that subgroup. However, if the D operator is inserted, the meaning of *vote to accept the proposal* shifts to a predicate that applies only to events whose agents consist of individuals x such that each x voted (i.e. cast a ballot) to accept the proposal. This shifted predicate satisfies the stratified-reference requirement of *all* and results in stronger truth conditions, requiring unanimity.

A related phenomenon, described in Kuhn (2014) based on examples by Benjamin Spector (p.c.), can be explained in a similar way:

- (41) a. All the axioms are consistent.
 b. #All the axioms are inconsistent.

In the absence of *all*, we expect both of these predicates to be in principle ambiguous between a collective and a distributive reading, though the distributive reading of (41b)

is implausible given the usual assumption that any axiom is internally consistent.¹⁰⁴ The collective reading of (41a) is available because it distributes down to subgroups: if a set of axioms is consistent, then any subset is too. As for (41b), it implausibly attributes the property of being inconsistent to each individual axiom. This suggests that its only reading is distributive. Since this reading entails that each axiom is internally inconsistent, this reading is implausible, and the sentence is deviant as a result. The collective reading of (41b) is correctly predicted to be unavailable because the collective interpretation of *be inconsistent* does not distribute down to subgroups (Jeremy Kuhn, p.c.): as shown by the example $\{a < b, b < c, c < a\}$ where $<$ is taken to be a total order, a set of axioms can be inconsistent when taken together even if any of its proper subsets is internally consistent. As for the distributive reading of (41b), it is correctly predicted as a result of applying the D operator to the predicate.

Since the examples in (42) are very similar to the minimal pair in (41), it might at first sight seem surprising that they are both acceptable (Kuhn 2014):

- (42) a. All the students agreed about what book to read.
 b. All the students disagreed about what book to read.

However, this is predicted by the present account. The key observation, as Kuhn points out, is that whenever a group disagrees, at least two different opinions are represented among its members. This always makes it possible to divide any group of disagreeing students into possibly overlapping subgroups such that each subgroup has at least two opinions represented in it. This means that just like *be consistent* and *agree* but unlike *be inconsistent*, *disagree* has subgroup distributivity. In other words, each disagreeing group can be divided into small subgroups such that each subgroup internally disagrees. Therefore, no application of a distributivity operator is required in order to satisfy the presupposition of *all* in (42b).

The D operator is not the only distributivity operator that can act as a repair strategy. As shown in the Appendix, the output of the distributivity operator has stratified reference with respect to any downward monotonic granularity level that is at least as coarse as the granularity level of the operator. Chapter 8 has proposed, following Schwarzschild (1996), that when the context provides a set of salient nonatomic entities, the nonatomic distributivity operator Part becomes available. Since *all* sets its granularity parameter to the downward monotonic predicate $\varepsilon(*\text{agent}(e))$ and since

¹⁰⁴ The collective reading of (41a) is stronger than its distributive reading (imagine a situation where each axiom is internally consistent but one of them contradicts the others). That both readings are available can be shown by embedding (41a) into a downward-entailing context. The following sentence can easily be imagined to be true on the collective reading but false on the distributive reading:

(i) If all the axioms in your solution are consistent, you will get extra credit.

this predicate can apply to nonatomic entities as long as they are sufficiently small, we expect that *all* should be able to trigger intermediate readings via meaning shift. Example (43), discussed by Brisson (2003: 173) and attributed to Veneeta Dayal (p.c.), shows that this prediction is correct. Ordinarily, the predicate *elect a president* lacks subgroup distributivity, and an analogous predicate is described in Taub (1989) as being incompatible with *all*. However, in the context of a yearly elementary school vote where each grade elects its own president, we can say the following:

(43) All the students elected a president.

In the given context, this sentence does not entail that there was one president that all the students elected, nor does it entail that each individual student elected a president. Thus, the relevant reading is intermediate, as expected from the insertion of the Part operator.

10.7 From *all* to *each*: different granularity levels

Let me now turn to the relationship between *all* and distributive quantifiers like *each* and *every*. As we have seen earlier, *numerous*-type collective predicates are incompatible with all of these items, while *gather*-type predicates are incompatible with *each* but compatible with *all*. In the previous section, I suggested that this is because *all* requires the verbal predicate to apply to subevents whose agents are small in number. This means that we can think of *all* as a sieve. In order for a predicate to pass through this sieve, it must distribute down to sufficiently small parts. In fact, *all* is like a coarse sieve, in the sense that the parts are allowed to have nonatomic agents. I suggest that we can think of *each* as a fine sieve. Like *all*, it imposes distributivity, but unlike *all*, it imposes distributivity down to atomic individuals. A predicate like *smile* will pass through both sieves because it distributes down to atomic agents, while a predicate like *gather* passes only through the coarse sieve because it distributes only down to subgroups.

More formally, to explain why *gather* distinguishes between *each* and *all*, I assume, following Kuhn (2014) and previous chapters, that *each* distributes over events with atomic individuals while *all* distributes over events whose agents must be small in number but need not be atomic. The entry from Section 9.5, repeated here, implements this assumption:

$$(44) \llbracket \text{each} \rrbracket_{\text{determiner}} = \lambda P \lambda \theta \lambda V \lambda e. \theta(e) = \bigoplus P \wedge \llbracket D\theta \rrbracket(V)(e) \\ = \lambda P \lambda \theta \lambda V \lambda e. \theta(e) = \bigoplus P \wedge e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$$

On this view, determiner *each* divides the relevant event into parts which satisfy the verb phrase and whose agents are atoms. For a sentence like *each boy smiled*, this entry yields the following meaning:

$$(45) \quad \exists e. *agent(e) = \bigoplus \text{boy} \wedge e \in * \lambda e' \left(\begin{array}{c} *smile(e') \wedge \\ Atom(agent(e')) \end{array} \right)$$

This says that there is an event e whose agent is all the boys and which consists entirely of smiling events whose agents are atoms. Since *agent* is a sum homomorphism, we know that these atoms must sum up to $\bigoplus \text{boy}$. Since every boy is an atom (see Section 2.6.1), this entails that the agents of the smiling events are boys.

Since any distributive predicate has stratified reference along the *agent* dimension with granularity *Atom*, it will be compatible with *each*. Collective predicates do not have this property, and will therefore generally lead to category mistakes. For example, the following sentences would require the existence of gathering events and being-numerous events whose agents are individual boys.

$$(46) \quad \llbracket *Each \text{ boy gathered} \rrbracket \\ = \exists e. *agent(e) = \bigoplus \text{boy} \wedge e \in * \lambda e' \left(\begin{array}{c} *gather(e') \wedge \\ Atom(agent(e')) \end{array} \right)$$

$$(47) \quad \llbracket *Each \text{ boy was numerous} \rrbracket \\ = \exists e. *agent(e) = \bigoplus \text{boy} \wedge e \in * \lambda e' \left(\begin{array}{c} numerous(e') \wedge \\ Atom(agent(e')) \end{array} \right)$$

In the present account, nothing hinges on the fact that *every* and *each* are syntactically singular while *all* triggers plural agreement in the count domain. This is in contrast with theories such as Winter (2001), in which *all* is the plural counterpart of *every* and its meaning is obtained by applying a type-shifter to *every*. One drawback of this view is that it does not straightforwardly extend to the mass domain. As we have seen in Section 10.5, when *all* combines with mass terms it triggers singular agreement, but it still imposes the same semantic constraints as in the count domain.

Given the similarities of *each* and *all*, the question arises whether both could be given the same lexical entry, differing only in their granularity parameter. One way to achieve this would be to adjust the entry for *each* in (44), which was derived in Section 9.3 from my treatment of the D operator, so that it becomes more similar to the presuppositional entry I have proposed for *all* in (28). Alternatively, the entry for *all* could be changed to match the operator-like entry for *each*. The first strategy would amount to giving *each* a similar treatment to the presuppositional entry proposed for it in Section 4.7, on which the entry for *all* in this chapter is modeled. This is possible, though as discussed in Section 9.3 there is a tradeoff between the presuppositional and the operator-like treatment of *each*. The former captures the parallels between adverbial *each*, *for*, and *of*, while the latter captures the parallel between *each* and the D operator. The second strategy would amount to giving *all* an operator-like treatment, modeled on the operator-like entry for *each* but with the *Atom* parameter replaced by $\varepsilon(*agent(e))$:

(48) Hypothetical alternative entry for determiner *all*:

$$\lambda P \lambda \theta \lambda V \lambda e. \theta(e) = \bigoplus P \wedge e \in {}^* \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \varepsilon({}^* \theta(e))(\theta(e')) \end{array} \right)$$

Instead of presupposing that the verbal predicate distributes down to events whose θ s are small in number as the entry in (44) does, the entry in (48) would simply distribute the predicate down to these events itself. I do not follow this strategy because, as we will see, it would lead to an overgeneration problem in connection with cumulative readings, to which I turn next.

10.8 *For* and *all* block cumulative readings

As we saw at the beginning of this chapter, the presence of *all* can block cumulative readings that would otherwise arise from the interplay of definite plurals and indefinite numerals (see Section 2.8). The contrast in (49) illustrates this behavior (Zweig 2008, 2009):

- (49) a. All the safari participants saw thirty zebras.
 Unavailable cumulative reading: Each safari participant saw at least one zebra, and thirty zebras were seen overall.
- b. Three safari participants saw thirty zebras.
 Available cumulative reading: Three safari participants each saw at least one zebra, and thirty zebras were seen overall.

It is surprising that (49a) does not have the cumulative reading of (49b), because this reading cannot be ruled out in terms of lack of plausibility. For example, suppose that (49b) is uttered in a context where there are only three safari participants. In this context, the noun phrases *three safari participants* and *all the safari participants* involve reference to the same plural individual. Yet, as (49a) shows, it is not possible to use them interchangeably. Only the former gives rise to a cumulative reading.

While the prototypical examples of cumulative readings involve discrete domains such as those in which count nouns denote, the concept is easily generalized to continuous domains such as time. In the following example, adapted from Brasoveanu (2013), a plural temporal entity stands in a cumulative relation with a plural sum of individuals.

- (50) This book is the product of five hundred hours of interviews with two hundred individuals.

In (49), we were able to relate each safari participant to a subset of the zebras; but in (50), we cannot relate each hour to a subset of the individuals, because some interviews may not last for an integer number of hours. Rather, the cumulative reading

of (50) expresses that there is a way to divide the five hundred hours of interviews into potentially overlapping parts so that each part can be related to at least one of the individuals in question. Thus, as we generalize cumulative readings to continuous domains, we need to move from distributivity over atoms to distributivity over parts of a cover (Schwarzschild 1996). In fact, cover readings also arise in discrete domains, where they have been previously assimilated to cumulative readings (Landman 1996). The following sentence can be true even if no fire fighter individually put out any single fire:

- (51) Four hundred fire fighters put out twenty fires.

Not all expressions that involve reference to time intervals are equally able to give rise to cumulative readings:

- (52) a. John saw thirty zebras.
b. John saw thirty zebras in an hour.
c. John saw thirty zebras for an hour.

Suppose John saw thirty zebras pass by him one at a time. This scenario can be felicitously described by (52a) and (52b) but not by (52c) (Krifka 1992, Eberle 1998). Now suppose John had an entire herd of thirty zebras in his field of view from noon to 1pm. Sentences (52a) and (52c) can felicitously describe this scenario, but sentence (52b) is not as natural.

The one-at-a-time scenario is analogous to situations that verify cumulative readings in the sense that each of the thirty zebras is matched to a time interval, and these time intervals sum up to a three-hour timespan (perhaps with some gaps in between the zebras). The simultaneous scenario is closer to a distributive reading in the sense that each of the zebras is matched to the whole timespan. In this sense, *in*-adverbials appear to be compatible with cumulative readings, while *for*-adverbials appear to impose stronger conditions that can block cumulative readings.

Let me now show how the same stratified-reference presupposition that rules out *numerous*-type collective predicates also explains the absence of cumulative readings in sentences like (49a). The denotation of the verb phrase *see thirty zebras* is as follows:

- (53) $\llbracket \text{see thirty zebras} \rrbracket = \lambda e. * \text{see}(e) \wedge |* \text{theme}(e)| = 30 \wedge * \text{zebra}(* \text{theme}(e))$
(True of any potentially plural seeing event whose themes sum up to thirty zebras.)

Given lexical cumulativity (see Section 2.7.2), among the events to which (53) applies there will be some which are composed of more than one seeing event. Thus if Tom, Dick, and Harry each saw a different sum of ten zebras, the sum of these three events is also a seeing event (call it e_0), and its theme is a sum of thirty zebras. Therefore the predicate (53) holds of e_0 .

Suppose now that *the safari participants* and therefore *all the safari participants* refers to the sum of Tom, Dick, and Harry. We can now explain why sentence (49a) does not have a cumulative reading. The lexical entry for *all* in (28) imposes the following stratified-reference presupposition on the verb phrase *see thirty zebras*, whose meaning is as in (53):

- (54) $SR_{*agent, \varepsilon}(*agent(e))(\llbracket \text{see thirty zebras} \rrbracket)$
 $= \forall e'. (*see(e') \wedge |*theme(e')| = 30 \wedge *zebra(*theme(e')) \rightarrow$
 $e' \in * \lambda e'' \left(\begin{array}{l} *see(e'') \wedge |*theme(e'')| = 30 \wedge *zebra(*theme(e'')) \wedge \\ \varepsilon(*agent(e))(*agent(e'')) \end{array} \right)$
 (True iff any potentially plural seeing event involving thirty zebras can be divided into (sums of) seeing events whose agents are each small in number compared to the agent of e , and which each have thirty zebras as their theme.)

This presupposition fails because the event e_o and any number of similar events falsify it. The event e_o is a plural seeing event whose themes sum up to thirty zebras. While it can be divided into seeing events whose agents are smaller in number than Tom, Dick, and Harry, none of these events has as many as thirty zebras as its theme.

One may think at first sight that the stratified-reference presupposition is too weak and that this account overgenerates by wrongly predicting sentence (49a) to be felicitous in a scenario in which there are many safari participants and thirty zebras, and subgroups of two or three safari participants saw these thirty zebras but no single individual saw every one of the thirty zebras. However, the existence of the event e_o suffices to make (54) fail even though this event is unrelated to anything that transpires in the scenario at hand. This is because (54) uses universal rather than restricted stratified reference and therefore quantifies over all seeing-thirty-zebras events whatsoever (and therefore also over any events in the model that are like e_o) rather than only those that are part of the scenario.¹⁰⁵

This account leads us to expect that *all* should not prevent cumulative readings across the board. For example, it should block them in connection with verb phrases like *see thirty zebras*, but not with verb phrases like *see zebras*:

- (55) All the safari participants saw zebras.

¹⁰⁵ My analysis in this section is essentially the same as in Champollion (2010b). Later, in Champollion (2015b), I analyzed *all* as well as *each*, *for*-adverbials, and pseudopartitives in terms of restricted stratified reference, in order to maintain a parallel analysis across domains. As discussed in Chapter 4, restricted stratified reference does not rule out predicates that fail to apply to events outside of the present scenario. This would lead to overgeneration in the present case. No corresponding problem seems to affect the analysis of *for*-adverbials and pseudopartitives in terms of restricted stratified reference. The analysis of those items in terms of restricted rather than universal stratified reference was adopted in Champollion (2015b) in response to arguments in Piñón (2015) and Schwarzschild (2015).

The relevant difference between *see thirty zebras* and *see zebras* is whether the object is bounded or unbounded. While this is intuitively similar to the difference between the telic predicate *eat thirty apples* and the atelic predicate *eat apples*, it might seem that this analogy breaks down because both *see zebras* and *see thirty zebras* are atelic. However, remember that unlike *for*-adverbials, *all* specifies the dimension parameter to a thematic role. Just as *eat apples* but not *eat thirty apples* distributes down the time dimension, *see zebras* but not *see thirty zebras* distributes down the agent dimension.

Specifically, the plural *zebras* in (55) is dependent on *all* and can indeed be analyzed essentially along the same lines as a cumulative reading. I assume that the literal meaning of *zebras* is *one or more zebras*, and that its full interpretation as *two or more zebras* is due to an implicature (see Spector 2007, Zweig 2008, 2009). This means that the verb phrase of (55) is literally interpreted as in (56):

- (56) $\llbracket \text{see zebras} \rrbracket = \lambda e. * \text{see}(e) \wedge * \text{zebra}(* \text{theme}(e))$
 (True of any potentially plural seeing event whose themes sum up to one or more zebras.)

Unlike *see thirty zebras*, this verb phrase satisfies the stratified-reference presupposition on the plausible assumption that any sum of seeing events whose themes are zebras consists of seeing events whose agents are atomic individuals and whose themes are zebras:

- (57) $\text{SR}_{\text{agent}, \varepsilon(\text{agent}(e))}(\llbracket \text{see zebras} \rrbracket)$
 $= \forall e'. (* \text{see}(e') \wedge * \text{zebra}(* \text{theme}(e'))) \rightarrow$
 $e' \in * \lambda e'' \left(\begin{array}{l} * \text{see}(e'') \wedge * \text{zebra}(* \text{theme}(e'')) \wedge \\ \varepsilon(* \text{agent}(e))(* \text{agent}(e'')) \end{array} \right)$
 (True iff any seeing event, or sum of seeing events, whose theme is a sum of zebras can be divided into (sums of) seeing events e'' such that the agent of any e'' is small in number compared to the agent of e , and the theme of e'' is one or more zebras.)

I have explained why the verb phrase *see thirty zebras* lacks a cumulative reading in a sentence like (49a). But what accounts for its distributive reading in that sentence? As discussed in Section 10.5, I assume that a covert distributivity operator can apply to the verb phrase before it combines with the subject. When this operator is coindexed with *agent* and applied to the verb phrase *see thirty zebras* as defined in (53), it yields the result in (58):

- (58) $\llbracket \text{D}_{\text{agent}} [\text{see thirty zebras}] \rrbracket$
 $= \lambda e. e \in * \lambda e' \left(\begin{array}{l} * \text{see}(e') \wedge \\ |* \text{theme}(e')| = 30 \wedge * \text{zebra}(* \text{theme}(e')) \wedge \\ \text{Atom}(\text{agent}(e')) \end{array} \right)$
 (True of any sum of potentially plural seeing events each of which has an atomic agent and a sum of thirty zebras as its theme.)

Unlike (53), this verb phrase satisfies the stratified-reference presupposition of *all*. More generally, the output of the distributivity operator in (31) always has stratified reference along the dimension specified by the role θ with which (31) is coindexed. The intuition is the following (see the Appendix for a formal proof). Suppose Tom, Dick, and Harry each saw thirty zebras. Then each of these three events is in the denotation of the nondistributive verb phrase (53). Since each of them has an atomic agent and since $P \subseteq *P$ for any P , each of these events is also in the denotation of the distributive verb phrase (58). Moreover, any sum of such events is also in the denotation of (58). In general, all events in the denotation of (58) will consist of such events, hence (58) has stratified reference. As we have already seen in Section 10.6, the distributivity operator can be seen as a repair strategy. When it applies to a verb phrase that would otherwise fail the presupposition of *all*, the result satisfies it.

At the beginning of this chapter, I mentioned that both *all* and *every/each* block cumulative readings, but only *all* licenses dependent plurals.

- (59) a. All the students had red noses.
b. #Every student had red noses.

As observed in Section 6.2, *for*-adverbials license dependent plurals as well:

- (60) John wore yellow neckties at night for a week.

These facts can be accounted for by building on the theory of dependent plurals in Zweig (2008, 2009). As discussed in Section 2.6.2, this theory essentially generates dependent plurals by combining the account of cumulative readings in Landman (1996) with an inclusive view of the plural, on which the *two or more* component is the result of an implicature (e.g. Spector 2007). The inability of *every* and *each* to license dependent plurals is explained by the assumption that they force the *two or more* component to be computed in their distributive scope. For example, *every student* in (59b) forces the *two or more* implicature to be computed within its scope, while *all the students* in (59a) imposes no such requirement.¹⁰⁶

Taken by itself, that theory predicts that all (and only) items that license dependent plurals can interact with items in their syntactic scope to give rise to cumulative readings. Since *all* and *for*-adverbials can license dependent plurals but not cumulative readings, this is the wrong prediction, as acknowledged by Zweig (2009: n. 38).

¹⁰⁶ As I mentioned in Section 8.6, the assumption that implicatures can and sometimes must be interpreted below the propositional level is independently motivated but not uncontroversial; see Schlenker (2016: sect. 22.2) for a discussion of relevant issues. Zweig motivates the fact that noun phrases headed by *every* and *each* always trap implicatures in their nuclear scope by assuming that these noun phrases always undergo quantifier raising. However, this assumption is not compatible with the account of *every* and *each* I have given in Section 9.5; and since it only applies to noun phrases, it leaves open why the plural object in *The students each had red noses* is not dependent on the subject. I stipulate that both adverbial *each* and distributive determiners trap implicatures in their scope, and I leave open how to independently motivate this assumption. For an expanded discussion of dependent plurals, see Champollion (2010b: ch. 9).

Stratified reference removes this problem by providing an independent factor that explains why *all* and *for*-adverbials block cumulative readings even though they license dependent plurals.

10.9 Open problems

Throughout this chapter, I have acknowledged a number of phenomena that the present account does not explain, such as the fact that some contexts improve the ability of certain *numerous*-type predicates to appear with *all* (Section 10.3), the existence of *numerous*-type predicates that might be argued to be subgroup-distributive (Section 10.4), and the differences in acceptability and interpretation that arise in connection with predicates like *build a raft* and *perform Hamlet* (Section 10.4 again). These predicates appear to tolerate collective readings with *all* for some speakers even if they do not license subgroup distributivity. As described in Section 10.4, this category also includes predicates modified by collectivizing adverbials like *build a raft together*. A possibly related phenomenon is the ability of adnominal *together* to prevent *all* from shifting collective predicates that lack subgroup distributivity into distributive predicates. The following examples all lose their collective interpretations and acquire distributive interpretations when the word *together* is removed:

- (61) Jacob's sons each became the head of a tribe, and all the tribes together were called Israel.¹⁰⁷
- (62) We gotta stick together: credit unions, leagues, and CUNA; all the elements together are unbeatable.¹⁰⁸
- (63) All the data together are pretty compelling.¹⁰⁹

As mentioned in Section 10.4, one possible explanation of such exceptions to the subgroup distributivity constraint is due to Brisson (2003), who assumes that the scope of the D operator can be optionally restricted to a silent event predicate DO. Brisson assumes that this predicate is present in sentences with activity and accomplishment predicates; for the stative examples (61) through (63), one could additionally assume that adnominal *together* inserts some predicate similar to DO. This kind of predicate effectually neutralizes the distributing effect of *all*, if we assume that its host noun phrase only takes scope over [*D_{agent}* DO] to the exclusion of the rest of the verb phrase.

- (64) [All the boys [*D_{agent}* DO]] [built a raft]. *collective*

¹⁰⁷ Attested example, accessed May 25, 2016. <http://twelvetribe.com/articles/why-twelve-tribes>.

¹⁰⁸ Attested example, accessed May 25, 2016. <https://www.ohiocreditunions.org/Advocacy/Pages/default.aspx>.

¹⁰⁹ Attested example, accessed May 25, 2016. <http://www.du.edu/give/faculty-linseman.html>.

However, I do not see a way to implement this approach without overgenerating cumulative readings in sentences involving *all* and activity predicates. Example (65) contains an activity predicate that contains a bounded argument, and it lacks a cumulative reading:

- (65) All the linguistics majors dated five chemistry majors.
(adapted from Zweig 2009)

The cumulative reading becomes available when adnominal *together* or *between them* is added:

- (66) All the linguistics majors { together / between them } dated five chemistry majors.

It appears that *together* and similar adverbials neutralize the stratified-reference constraint imposed by *all* and make collective and cumulative readings available that would otherwise violate it. Perhaps a covert version of such an adverbial is available in Dowty's dialect for predicates like *build a raft*.

I have not attempted to formalize the effect of such adverbials and how to constrain the distribution of their covert counterparts. One intuitive possibility is that *together* shifts the definite plural to which *all* attaches from a sum to an impure atom (see Section 2.8), and that impure atoms satisfy subgroup distributivity vacuously. However, making this precise is not straightforward, and I leave the problem unsolved here. More generally, it is an open question in what ways the interaction of adnominal and adverbial *together* with the semantics I have assigned to *all* can be derived from current theories about the meaning of *together* (see Lasersohn 1998a, Moltmann 2004).

Another problem is raised by the ability of *all* to license cumulative readings in passive sentences:

- (67) All the games were won by the Fijians and the Peruvians.

Here, a possible explanation is that the *by*-phrase is adjoined at sentence level and remains outside of the scope of *all the games*, and therefore outside of the stratified-reference presupposition.

Turning to *for*-adverbials, I have noted in Section 10.8 that cumulative readings are in general unavailable in their presence. However, a reviewer notes that a cumulative reading seems to be available in examples like the following:

- (68) a. The search committee interviewed six job candidates for three hours.
b. Mary baked eight pans of cookies for four hours.¹¹⁰

¹¹⁰ Adapted from an attested example, accessed May 25, 2016. <http://lulwanda2011.blogspot.com/2011/07/teacher-conference-ending.html>.

In a pilot study carried out via TurkTools (Erlewine & Kotek 2016), native speakers of American English rated (68a) significantly lower as an appropriate description of a scenario where the interviews take place in sequence than as a description of a scenario when they all take place simultaneously. However, the absolute ratings were high, and the same contrast could not be reproduced for (68b).¹¹¹

Since these types of examples violate stratified reference on their cumulative reading, there is no way to derive it on the present account. In fact, none of the accounts of *for*-adverbials reviewed in Chapters 5 and 6 fare any better, and I cannot identify any algebraic notion of unboundedness that would apply to the predicates in (68). As the reviewer notes, some sort of an intermediate account that allows for a constrained (perhaps context-determined) availability of cumulative readings is desirable. Whatever escape mechanism allows bounded predicates to combine with *for*-adverbials in examples like (68) can perhaps be related to whatever turns out to be the process that allows certain non-subgroup-distributive (and hence bounded) predicates to combine with *all* in examples like (64).

10.10 Summary

In this chapter, I have proposed that *all* is a distributive item. Like *each*, it requires the verbal predicate with which it combines to have stratified reference with respect to its own thematic role. Unlike *each*, *all* does not require distributivity down to atoms but merely down to small subgroups. I have suggested that those predicates with which *all* is compatible on a collective reading distribute down to subgroups of small cardinality. Informally, if a collective predicate holds of a plural entity *X*, then it will be compatible with *all* on a collective reading just in case there is a way to divide *X* into small, possibly overlapping subgroups (usually pairs or triples) such that the predicate applies to each of these subgroups.

The starting point for the present account was the intuition that *see zebras* is to *see thirty zebras* what *eat apples* is to *eat thirty apples*, and that setting distributive readings aside, *all* rules out the non-subgroup-distributive predicate *see thirty zebras* for the same reason as *for two hours* rules out the telic predicate *eat thirty apples*.

Modeling *all* as a distributive item provides us with an explanation of its puzzling scopal behavior and allows us to link it to *for*-adverbials and other distributive constructions. Both *all* and *for*-adverbials reject predicates that would give rise to cumulative readings. I have shown that these predicates do not fulfill the stratified-reference presuppositions imposed by *all* and by *for*. Predicates with bare plurals, like *see zebras*, satisfy stratified reference and give rise to cumulative readings.

The facts concerning cumulative readings have not often been discussed in previous work on *all* and *for*-adverbials. As far as I know, the present account is the first to

¹¹¹ I am grateful to Hanna Muller and Linmin Zhang for their help with this pilot study.

establish a formal connection between the scopal behavior of *all* and of *for*-adverbials. The literature on *all* has instead concentrated on why *all* is compatible with certain collective predicates like *gather*, but rejects others like *be numerous*. I have proposed an account of these two different types of collectivity in terms of stratified reference. *Numerous*-type predicates are incompatible both with *each* and with *all* because they lack stratified reference. This explains why *each* and *all* do not allow these predicates to have collective readings. *Gather*-type predicates violate the presupposition of *each*, but not of *all*, because they have stratified reference down to subgroups. This explains why *all*, but not *each*, allows them to have collective readings.

The D operator can rescue both *numerous*-type and *gather*-type predicates from presupposition failure. The result is a distributive reading, which is compatible with *all* as long as the subject of the sentence is a group noun.

Unlike accounts such as Dowty (1987) and Winter (2001), the present account does not build on the fact that the noun phrase *all the boys* is often understood as involving reference to more boys than the definite plural *the boys*. The behavior of *all* is not linked to these maximality effects. I have argued that this is a welcome fact because other quantifiers such as *most* do not involve maximality effects but pattern with *all* in other respects.

Conclusion

This chapter concludes the book by summarizing its main insights and results in a chapter-by-chapter summary (Section 11.1) and by offering some suggestions for further research (Section 11.2).

11.1 Chapter-by-chapter summary

In this book, I have developed a new approach to the semantics of distributivity, aspect, and measurement, three domains which are traditionally addressed by separate areas of research within formal semantics. By triangulating between these domains, I have arrived at a unifying perspective from which I made theoretical and empirical contributions to the study of the formal semantics of natural language.

My main theoretical contribution, introduced in **Chapter 1** and laid out throughout the book, is the notion of *stratified reference*, a concept that requires a predicate that applies to an entity—be it a substance, an event, or a plural individual—to also apply to the parts into which this entity can be decomposed along some dimension and down to some level of granularity. The concept is general enough to subsume a wide range of previous proposals, yet formally precise enough to make testable predictions and to transfer insights across traditional boundaries. The resulting framework, *strata theory*, is intended as a bridge that spans a number of semantic oppositions: singular/plural, count/mass, telic/atelic, and collective/distributive. While it has often been observed that these semantic oppositions are similar, and proposals have been made to bring some of them under the same umbrella, this work is the first one to propose a fully unified account. Intuitively, the concept that underlies each of these oppositions is the difference between boundedness and unboundedness. Singular, count, telic, and collective predicates are all delimited or bounded, in ways that set them apart from plural, mass, atelic, and distributive predicates. When it comes to formally describing what boundedness amounts to, characterizations in the semantic literature have tended to be limited to one domain: aspect, distributivity, or measurement only. Stratified reference provides a characterization that works in all of these domains. It builds on the same background assumptions as many previous theories and frameworks based

on classical extensional mereology and algebraic semantics (Lønning 1987, Link 1998a, Krifka 1998, Landman 2000).

These theories and assumptions are presented in explicit and distilled form in **Chapter 2**, with a focus on areas in which no consensus has been reached, such as the meaning of the plural morpheme, the question whether the meanings of verbs are inherently pluralized, the formal properties of thematic roles, and the compositional process. This chapter is intended as a reference point for future researchers and as an introduction to the relevant parts of the formal semantic literature.

My main empirical contribution is the observation that a large class of nominal and verbal constructions impose analogous unboundedness constraints on a predicate denoted by one of their constituents. A representative selection of what I have called *distributive constructions*—*for*-adverbials, pseudopartitives, and adverbial *each*—is described in **Chapter 3** and onwards. The chapter includes simplified Logical Forms for these constructions that provide a scaffold on which the theory in the rest of the book is built. Distributive constructions give us an empirical handle on the conceptual question of how to characterize unboundedness. For example, the fact that *for an hour* can modify the unbounded predicate *eat apples* but not the bounded predicate *eat thirty apples* makes it possible to constrain the space of options for formal definitions of unboundedness by studying the algebraic properties of these and related predicates (Krifka 1998). Stratified reference emerges from a systematic investigation of these constructions and of previous theories that account for their behavior within the framework of algebraic semantics.

Since unboundedness is a property of predicates, and since predicates are properties, it is natural to think of unboundedness as a higher-order property. Indeed, previous work in algebraic semantics has used higher-order properties such as cumulative or divisive reference to characterize different facets of unboundedness (e.g. Link 1998a, Krifka 1998). Such properties are a useful stepping stone towards a formal characterization of unboundedness, but they are too rigid to provide a nuanced understanding of the differences between these facets. For example, distributivity and atelicity can both be seen as facets of unboundedness; but one and the same predicate can be atelic without being distributive or vice versa, or distributive with respect to one thematic role but not another. **Chapter 4** presents stratified reference as a formalization of unboundedness and as a means to capture the parallels between the semantic oppositions in a uniform way. After giving a brief overview over the empirical phenomena that have been discussed under the rubric of distributivity, the notion of stratified reference is gradually developed as a generalized notion of distributivity. It is then used to formulate constraints that capture the behavior of distributive constructions and meaning postulates that predict distributive entailments of lexical predicates.

Two factors make it possible to identify a single formal property that describes unboundedness in all its facets. The first factor consists in using the same descriptive

terms for constituents that behave analogously across syntactically and semantically distinct distributive constructions. The terms *Key* and *Share* from the literature on distributivity turn out useful for this purpose, as does the newly coined term *Map*. The second factor is the combination of a common approach in semantics—using higher-order properties—with a common approach in syntax—using parameters. This leads to the conceptualization of stratified reference as a *parametrized higher-order property*. Stratified reference builds on the basic intuition behind algebraic semantic accounts, namely that atelicity, distributivity, and related concepts can be defined in terms of a predicate applying to the parts of an event or entity, and generalizes it by adding parameters that allow us to explicitly model varying dimensions and granularities. These parameters turn out to provide an appropriate middle ground between rigidity and flexibility that captures the ways in which distributive constructions differ without losing track of their common core. Following Piñón (2015) and Schwarzschild (2015), I have taken a further step away from ordinary higher-order properties by restricting stratified reference to the parts of a single entity x , rather than requiring it to apply to all entities to which the predicate applies. In effect, these moves make stratified reference into a relation that is higher-order on its predicate argument and on its two parameters and first-order on its entity argument. However, since its purpose is still conceptually close to higher-order properties, I will refer to it as the property-based perspective on stratified reference:

(1) **Property-based perspective**

$$\text{StratifiedReference}_{\text{dimension, granularity}}(\text{Predicate})(x) \stackrel{\text{def}}{=} x \in {}^*\lambda y. \left(\begin{array}{l} \text{Predicate}(y) \wedge \\ \text{granularity}(\text{dimension}(y)) \end{array} \right)$$

The *granularity parameter* allows us to model the varying amounts to which distributivity will reach down to subparts in distributive constructions: to atoms or small subgroups in some cases, and to contextually salient levels of granularity in others. This parameter is motivated in part by the need to account for the minimal-parts problem, as was done in **Chapter 5**. This problem arises from the fact that some eventualities and substances fail to distribute at very small scales because they have parts that are too small to satisfy certain mass terms and atelic predicates. This is a challenge for characterizations of atelicity that look at all smaller events (as in the case of divisive reference, Krifka 1998) or intervals (as in the case of the subinterval property, Dowty 1979). Nondivisive atelic predicates such as *waltz* and *pass on from generation to generation* make it necessary to relativize these concepts, for example by equipping them with a minimal-length threshold so that they ignore what happens at very short intervals below this threshold.

By making a virtue out of necessity and elevating this threshold to a central part of the theory—the granularity parameter—it becomes possible to avoid the minimal-parts problem. Different settings of the parameter lead to nuanced predictions

concerning the interaction between the predicate and the length of the interval denoted by the complement of *for*. By varying the parameter, we may use stratified reference both to describe the length of the smallest events that count as waltzing or passing on from generation to generation, and to describe the requirements that *for*-adverbials impose on the properties they modify. I did not fully recognize these two tasks as conceptually distinct until the response articles to Champollion (2015c), particularly Piñón (2015) and Schwarzschild (2015), helped me realize it. If a *for*-adverbial is like a sieve and the events in the denotation of the predicate it modifies are like grains of sand, the first task amounts to describing the size of the grains, and the second amounts to describing the size of the holes in the sieve (Champollion 2015b). As a part of the description of the constraint imposed by *for*-adverbials, stratified reference describes the size of the holes. This is what Chapter 5 focuses on. As a component of meaning postulates that describe what we know about predicates, stratified reference can describe the size of the grains that pass through the sieve. Some predicates like *waltz* will be fine-grained; other predicates like *pass on from generation to generation* will be more coarse-grained.

The *dimension parameter* captures the view that unboundedness may occur in time, in space, or along a measure function or a thematic role. Because of its traditional focus on cross-domain generalizations, mereology-based algebraic semantics lends itself well to a formal implementation of this view. In particular, various functions can be treated as of one and the same kind: thematic roles such as *agent* and *theme*, measure functions such as *temperature* and *volume*, and event properties such as *runtime* and *spatial extent*. As we have seen in Chapter 6, the latter parallel makes it straightforward to account for analogies between temporal measure adverbials (such as *run for an hour* vs. **run all the way to the store for an hour*) and spatial measure adverbials (such as *meander for a mile* vs. **end for a mile*) previously noted by Moltmann (1991) and Gawron (2009). More generally, the dimension parameter captures the fact that a distributive construction will typically impose only one kind of unboundedness at a time. For example, the fact that temporally unbounded predicates can be modified by temporal *for*-adverbials even when they contain a spatially bounded constituent (as in *flow from the jar to the floor for ten minutes*) is unsurprising on this view, and differences in interpretation between temporal and spatial *for*-adverbials (as in *push carts all the way to the store for fifty minutes* versus *for fifty meters*) find a natural explanation.

Theories of aspect are typically not designed as ways to explain what is wrong with measure constructions like **three degrees of water* or **three pounds of book*. However, relevant connections have occasionally been noted (Krifka 1998, Schwarzschild 2006). The parallel becomes intuitive once we think of the verb phrase *run for three hours* in connection with the pseudopartitive *three hours of running*. Chapter 7 has exploited the formal parallel between the domains of aspect and measurement developed in Chapter 4 to explain the linguistic relevance of the difference between intensive

measure functions like *temperature* and extensive ones like *runtime*. Treating the two constructions as semantically equivalent made it possible to push the limits of theories designed for only one of the two domains to which they are traditionally seen as belonging. Stratified reference correctly predicts that distributive constructions disallow measure functions that generally return the same value on an entity and on its parts. For example, just as *run for three hours* requires *run* to apply to temporally shorter parts of the event to which *run* applies, **three pounds of book* would require *book* to apply to lighter parts of the entity to which *book* applies, and **three degrees Celsius of water* would require the existence of colder parts of the entity to which *water* applies. The fact that stratified reference relativizes unboundedness to just one dimension or measure function at a time made it possible to subsume the insight of Schwarzschild (2006), and to account for examples like *five feet of snow* in spite of the fact that not every part of a five-foot snow layer is less than five feet in height.

Throughout this book, I have used stratified reference for various purposes: to characterize the distributivity constraint in those constructions that impose it; to specify meaning postulates for words that exhibit distributivity down to various levels of granularity; and as a formalization of atomic and nonatomic distributivity operators. Starting in Chapter 8, I shifted from viewing stratified reference as a parametrized higher-order property to viewing it as a parametrized unary distributivity operator on predicates:

(2) **Operator-based perspective**

$$\text{StratifiedReference}_{\text{dimension, granularity}}(\text{Predicate}) \stackrel{\text{def}}{=} \lambda x. x \in * \lambda y. \left(\begin{array}{c} \text{Predicate}(y) \wedge \\ \text{granularity}(\text{dimension}(y)) \end{array} \right)$$

Since the two definitions sketched in (1) and (2) are equivalent, the move from the property-based perspective to the operator-based perspective is largely conceptual. The main difference results from whether stratified reference is implemented as a presuppositional requirement or as a predicate modifier. In both cases, the dimension and granularity parameters can be instantiated in whatever ways may be appropriate for different constructions and theoretical assumptions. Varying the value of the dimension parameter amounts to distributing over various thematic roles and spatiotemporal dimensions. Varying the value of the granularity parameter amounts to choosing between distributing over atomic entities like singular individuals and nonatomic entities like pluralities and temporal intervals. Chapter 8 exploited the operator-based perspective to synthesize and expand previous accounts of how verb phrases such as *build a raft* optionally acquire a distributive interpretation by covert distributivity operators. In particular, the differences between the atomic operator in Link (1987b), Roberts (1987), the nonatomic operator in Schwarzschild (1996), and their generalizations in Lasersohn (1998b), can be modeled and clarified by shifting

the values of the dimension and granularity parameters. Furthermore, setting the dimension parameter to *runtime* made it possible to transfer the notion of covert distributivity as a verb phrase shifter into the temporal domain, and making the granularity parameter anaphoric to a salient predicate helped export Schwarzschild's claim that nonatomic distributivity requires salient covers to that domain. This resulted in a new perspective on the puzzling scopal behavior of indefinites and numerals in the scope of *for*-adverbials, including the fact that indefinites in the syntactic scope of *for*-adverbials tend not to covary with these adverbials (*?John found a flea on his dog for a month*, Zucchi & White 2001).

The operator-based perspective on stratified reference naturally led to postulating a formal connection between covert and overt distributivity. This made it possible in **Chapter 9** to analyze distance-distributive items across languages as overt versions of distributivity operators, as suggested by Link (1991b) for the case of *each*. The granularity parameter, along with the notion of an anaphoric cover from Schwarzschild (2006), made it possible to account for the crosslinguistic variation between those distance-distributive items that only exhibit atomic distributivity, such as adverbial and adnominal *each* in English, and those that also distribute over salient nonatomic entities such as time intervals, such as adverbial and adnominal *jeweils* in German. Essentially, these two items were treated as overt versions of the atomic and the nonatomic distributivity operator respectively. The typological correlation between atomic distributivity and ability to be used as a distributive determiner observed by Zimmermann (2002b) turned out to be expected once the operator-based perspective was extended to distributive determiners such as *each* and *every*. Because stratified reference always provides access to the sum event and not just to its parts, this extension immediately explained why these determiners can participate in nondistributive phenomena with items outside of their syntactic scope, such as cumulative readings and nondistributive adverbial modifiers (Schein 1993, Kratzer 2000, Champollion 2010a).

The property-based and the operator-based perspective on stratified reference, as well as both the dimension and the granularity parameter, all come into play in **Chapter 10**, whose main focus is on explaining the behavior of *all* with respect to different collective predicates, such as *all the students gathered* versus **all the students were numerous* (Dowty 1987, Winter 2001). The impetus for this chapter came from the startling observation by Zweig (2009) that a noun phrase headed by *all*, such as *all the safari participants*, can lead to a cumulative reading when it combines with a verb phrase that contains an unbounded argument, such as the dependent plural in *saw zebras*, but not when the verb phrase contains a bounded argument, such as *saw thirty zebras*. The search for a formal property that sets these two predicates apart was facilitated by the property-based perspective. My guiding intuition was that *see zebras* is to *see thirty zebras* what *eat apples* is to *eat thirty apples*; however, the second pair captures the telic/atelic opposition while both expressions in the first pair are

atelic. Stratified reference allows us to model this situation as a difference in settings of the dimension parameter. *Eat apples* but not *eat thirty apples* distributes down the time dimension; *see zebras* but not *see thirty zebras* distributes down the agent dimension. Stratified reference instantiated with *time* is atelicity; stratified reference instantiated with *agent* is distributivity. If *for*-adverbials test for atelicity, *all* tests for distributivity.

While a traditional view on distributivity might lead us to expect that *all* is synonymous with other distributive items such as *each*, the fact that different items set the granularity parameter to different values leads us to expect otherwise. The fact that *all* but not *each* is compatible with collective predicates that exhibit subgroup distributivity finds a natural explanation in the assumption that *all* is a coarser sieve than *each*, in line with the characterization of *gather*-type predicates as subgroup distributive (Dobrovie-Sorin 2014, Kuhn 2014). Stratified reference was also used to formulate meaning postulates that capture the fact that *gather*-type predicates give rise to distributive inferences to subgroups. Finally, the distributive operator from Chapter 8 helped account for cases in which *all* appears to take away the collective interpretation of a predicate that can normally be interpreted either distributively or collectively.

11.2 Future work

This section, adapted in part from Champollion (2015b, 2015c), sketches some broader implications of strata theory and connections to other domains of linguistics.

Any theory of distributive constructions needs to specify the constraint that these constructions impose on their constituents (the nature of the sieve) as well as the reason these constituents satisfy it (the nature of the grains that pass through the sieve). In the case of one-word predicates, I have used stratified reference to formulate meaning postulates that describe their grain size. As we have seen in Chapter 8, complex predicates can also be characterized with respect to whether or not they satisfy stratified reference. A full account of aspect and distributivity in these cases will need to be complemented by a theory of how a given complex predicate ends up having or not having stratified reference. Certain overt modifiers, such as adverbial *each* and *together*, can determine whether the predicate that they modify is understood distributively or collectively. The question of how complex predicates end up being collective or distributive is analogous to the question of how complex predicates end up being atelic or telic, a process also known as aspectual composition (e.g. Krifka 1998). Stratified reference allows us to think about the effect of *each*, *together*, and distributivity operators and about aspectual composition as two sides of the same coin.

This also means that we can link problems that affect accounts of these processes. For example, Doetjes (2015) correctly notes that stratified reference does not rule out

incremental-theme verbs whose themes are downward-entailing modified numerals, such as (3a). That modified numerals pose problems for algebraic accounts of aspectual composition has been noticed many times (Egg 1994, Eberle 1998, Naumann 1999). Stratified reference has this problem in common with the subinterval property, which it is meant to generalize. Cumulative reference fares no better because an analogous problem occurs with upward-entailing modified numerals, as in (3b). Doetjes therefore proposes combining stratified reference with cumulative reference, following Landman & Rothstein (2012b). This will rule out both types of examples as desired, but unfortunately not (3c), discussed by Zucchi & White (2001). Likewise, the contrast in (3b), discussed by Mittwoch (1982), remains unexplained.

- (3) a. #He drank at most thirty glasses of water for three hours.
b. #He finished at least three books for three hours.
c. #John drank { some / a quantity of } milk for an hour.
d. John { ate / #ate something } for an hour.

This kind of behavior is puzzling for most if not all algebraic theories of aspect, including strata theory. The noun phrases that cause the sentences in (3) to sound odd seem to behave for the purposes of these theories as if they were quantized, at least along the relevant (temporal) dimension. An early feature-based theory of aspectual composition, Verkuyl (1972), grouped modified and unmodified numerals together by assigning both of them a [+SPECIFIED QUANTITY] feature, while bare plurals and mass nouns carried a [-SPECIFIED QUANTITY] feature. Algebraic notions like quantization, stratified reference, and the subinterval property are meant to make such features superfluous. But in noun phrases like those in (3), the effects of these two systems come apart (Verkuyl 2005: n. 3). A similar issue is discussed by Schwarzschild (2015) in connection with the word *line*. Other problematic predicates include *twig*, *rock*, and *sequence*. A helpful but ultimately inconclusive discussion of possible ways to address this problem is found in Zucchi & White (2001). Similarly, as discussed in Chapter 10, a number of collective predicates that are incompatible with *all* but that are still subgroup distributive, such as *be a group of less than five*, would be expected to be compatible with *all* under the account I have discussed here (Kuhn 2014). Finally, the constraint against cumulative readings of *all* described in Chapter 10 also rules out a cumulative reading when the verb phrase contains a delimited but nonquantized object (*All the linguistics majors dated several chemistry majors*, Zweig 2009). If a solution to these problems emerges in one domain, we may well be able to adapt it to the other domain.

If this book is on the right track, distributivity is ubiquitous. We just need to recognize it when it presents itself in unusual ways. I have made the case for this idea using *each*, *all*, *for*-adverbials, and pseudopartitives. Now that we know what we are looking for, it should be easy to find more distributive constructions. Here are some possible places to look:

German and Japanese split-quantifier constructions, in which a quantifier appears in adverbial position apart from the noun phrase over which it quantifies, are similar to adverbial-*each* distributive constructions in that they are incompatible with collective interpretations, and they are similar to pseudopartitive constructions in that their measure functions are subject to the same monotonicity constraint (Nakanishi 2004).

As discussed in Ursini (2006), directional prepositional phrases can be modified by measure phrases when they are unbounded (*three miles towards the beach*), but not when they are bounded (**three miles to the beach*). This points towards the possibility that this is a distributive construction. The measure phrase in these examples might be a Key, and its directional prepositional phrase a Share.

For-adverbials are not the only examples of aspectually sensitive constructions. As noted in Karttunen (1974) and Hitzeman (1991, 1997), *until* is also sensitive to the atelic/telic distinction. The same appears to be true for *since*, though the situation is more complicated here. In English, *since* requires the Perfect, which is often analyzed as introducing an Extended Now interval (Dowty 1979, von Stechow 2002a). This muddles the picture, but once we move to German, where the equivalent *seit* does not require the Perfect, we see the correlation emerge:

An Extended Now Perfect modified by *since* α may embed any aktionsart. German perfects modified by *seit* α may have these readings, though they are a bit marked. In contrast to English, *seit* α may combine with simple tenses as well, but then it behaves differently. The aktionsart modified must be a state or an activity. (von Stechow 2002a: 394)

The theory of the behavior of indefinites in the scope of *for*-adverbials presented in Chapter 8 can be extended to other modifiers that do not or not easily induce covariation of indefinites in their scope. In particular, habitual or generic sentences show analogous scopal effects to *for*-adverbials (Carlson 1977, Kratzer 2007). This is illustrated in the examples in (4), taken from Krifka, Pelletier, Carlson, ter Meulen, Chierchia & Link (1995: 39f.).

- (4) a. Mary smokes { cigarettes / *a cigarette }.
- b. Mary smokes { cigarettes / a cigarette } after dinner.

Just like in the case of *for*-adverbials, singular indefinites can covary with habitual operators when a salient level of granularity is provided (see also Rimell 2004).

- (5) a. Yesterday, Mary smoked { cigarettes / *a cigarette } for an hour.
- b. Last month, Mary smoked { cigarettes / a cigarette } after dinner for a week.

This fact suggests that the generic quantifier might carry a stratified-reference presupposition, and that it might be appropriate to fold strata theory into a more general theory of imperfective and generic/habitual sentences such as the one proposed in Deo (2009) for English and Gujarati and extended to *for*-adverbials in Deo & Piñango

(2011). Similar effects to the ones in (5) hold in Hindi (Ashwini Deo, p.c.), which is close to Gujarati. For more discussion and for a synthesis of Deo & Piñango (2011) and the present account, see Champollion (2013).

Other potential applications can be found in morphosyntax. Strata theory may help explain how boundedness is marked by semantic case in Finnish (Krifka 1992, Kiparsky 1998), by perfective prefixes in Slavic (Filip 2000), and by accusative adverbials in Korean (Wechsler & Lee 1996). Throughout this book, I have assumed that singular count nouns are interpreted as involving reference to singular entities but not sums. This was necessary in order to explain the contrast between *five pounds of books* and **five pounds of book*, and it is justified in English by the corresponding contrast in numeral phrases (*five books* vs. **five book*). Other languages, like Hungarian and Turkish, require nouns to be morphologically singular when they combine with numerals, and also when they are used as substance nouns in pseudopartitives. From the point of view of the present theory, this leads to the view that singular nouns in these languages and constructions can be interpreted as involving reference to sums. Theories that adopt this view (Farkas & de Swart 2010, Bale, Gagnon & Khanjian 2011) are compatible with the view developed here. This may be seen as an advantage for them over theories that reject this assumption (Ionin & Matushansky 2006).

I have focused on pseudopartitives like *three liters of water*. As noted by Schwarzschild (2002, 2006), true partitives like *three liters of the water* and comparatives like *more water* are subject to the same constraint on measure functions as pseudopartitives. An extension of the present account to true partitives is straightforward if we assume that the constituent *of the water* has divisive reference, stratified reference, or whatever is the relevant property of the substance nominal of pseudopartitives. However, the assumption that the *of*-PP has divisive reference is not uncontroversial: Ladusaw (1982), and many accounts that follow him, adopts it, but Matthewson (2001) argues against it.

While I have shown that the behavior of a large number of constructions can be reduced to one principle (namely, sensitivity to stratified reference), I have not addressed the question why this principle exists and why these constructions are sensitive to it. In formal semantics, this is not the kind of question that is typically answered, or perhaps even answerable. There is no agreement on whether it even needs to be asked. On the one hand, for the purposes of comparing formal semantic theories to each other, formal semantics usually pays attention to something similar to Chomskyan explanatory adequacy: “If a number of highly complex and apparently unrelated facts are reducible to a few simple principles, then these principles explain these facts” (von Stechow 1984a). On the other hand, we need not confine ourselves in this way: “we can seek a level of explanation deeper than explanatory adequacy, asking not only what the properties of language are but also why they are that way” (Chomsky 2001).

I do not know why there should be any constructions in language, let alone so many of them, that are sensitive to stratified reference or to the various properties it

captures. To answer this question, it may be worth looking for explanations in domains other than formal semantics, such as first-language acquisition. Stratified reference may conceivably help first-language learners distinguish the functions of different constructions. For example, learners must distinguish constructions that specify the quantity of a substance or event, such as pseudopartitives, from superficially similar constructions that specify non-quantity-related properties, such as attributive constructions (*three-pound strawberries*). Attributive constructions do not impose stratified reference and are therefore compatible with intensive measure functions, as illustrated by *three-degree water* (Schwarzschild 2006). Apart from sometimes misinterpreting the number word in pseudopartitives as referring to cardinality of a relevant set of objects, four-year-olds tend to correctly distinguish pseudopartitives from attributives (Syrett 2013). Similarly, various studies have suggested that children are sensitive to the atelic/telic opposition as early as three years old, raising the question of how much of it is innately specified (Crain 2011). If something like the boundedness/unboundedness opposition is among the building blocks of the language faculty, then we might expect that children access it early on, and possibly that a child will learn different constructions that involve this building block at the same age.

Another kind of explanation, as well as another avenue for further research, may be found in linguistic theories that study conceptual linguistic knowledge and the mental patterns and representations in which it is organized, such as cognitive semantics (Talmy 2011) and conceptual semantics (Jackendoff 1996). The metaphor I have used to explain stratified reference—namely that individuals, substances, and events occupy regions in an abstract space whose dimensions include thematic roles and measure functions as well as spatial and temporal dimensions—is reminiscent of the theory of conceptual spaces in Gärdenfors (2007). The words that introduce stratified-reference constraints, such as *for*, *until*, *of*, *each*, and *all*, belong to closed-class categories such as prepositions and determiners. Cognitive semantics has found that closed-class categories are highly constrained in the range of conceptual categories they can express. The relevant conceptual category in this case would be boundedness. While cognitive semantics is sometimes seen as opposed to formal semantics, this does not have to be so (Krifka 1998, Zwarts & Verkuyl 1994). We can make use of formal semantic techniques such as the ones I have developed here, and assume that expressions are interpreted by elements of conceptual structures rather than entities in the real world. The present system may then be seen as a step towards a model-theoretic characterization of such frameworks.

Appendix

Distributivity operators as repair strategies

At various places in this book, I have assumed that the distributivity operators D and $Part$ can act as repair strategies that shift the meanings of predicates so that they satisfy constraints imposed by distributive constructions. I have assumed that this is possible because the output of a distributivity operator has stratified reference with respect to any granularity level that is at least as coarse as the granularity parameter of the operator.

Here I prove a theorem that justifies this assumption. The guiding intuition behind the proof is based on the idea that the granularity parameter of a distributive construction acts as a threshold on the thickness of the strata to which the Share predicate is required to apply. Distributivity operators ensure stratified reference by applying the predicates that they modify to entities that are small enough to satisfy these thresholds. Therefore the output of the distributivity operator will satisfy stratified reference whenever its own granularity setting is fine enough. To put it differently, we may think of a distributive construction as representing a sieve, while a distributivity operator produces sand to be sent through that sieve. If C is the granularity level of the operator, C' that of the construction, and both C and C' have divisive reference, then the sand will be fine enough to pass through the sieve whenever $C \subseteq C'$.

A sieve that is coarse enough for a grain of sand to pass through should of course also be coarse enough for any part of that grain to pass through. Given this kind of reasoning, it makes sense to instantiate the granularity parameter on the distributive construction with a predicate that has divisive reference (see Section 2.3.5). I have implemented this assumption in all distributive constructions (see Section 4.6).

The following theorem refers the D and $Part$ operators repeated here from Chapter 9, and to the definitions of universal and restricted stratified reference repeated here from Chapter 4:

(1) **Definition: Event-based D operator**

$$\llbracket D_\theta \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ \text{Atom}(\theta(e')) \end{array} \right)$$

(Takes an event predicate V and returns a predicate that holds of any event e which can be divided into events that are in V and whose θ s are atomic.)

(2) **Definition: Event-based $Part$ operator**

$$\llbracket Part_{\theta, C} \rrbracket \stackrel{\text{def}}{=} \lambda V \lambda e. e \in * \lambda e' \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right)$$

(Takes an event predicate V and returns a predicate that holds of any event e which can be divided into events that are in V and whose θ s satisfy the contextually salient predicate C .)

(3) **Definition: Universal stratified reference**

$$\text{SR}_{d,g}(P) \stackrel{\text{def}}{=} \forall x \left[P(x) \rightarrow x \in *_{\lambda y} \left(\begin{array}{c} P(y) \wedge \\ g(d(y)) \end{array} \right) \right]$$

(P has universal stratified reference along dimension d with granularity g iff any x in P can be divided into one or more parts in P that are each mapped by d to something in g .)

(4) **Definition: Restricted stratified reference**

$$\text{SR}_{d,g}(P)(x) \stackrel{\text{def}}{=} x \in *_{\lambda y} \left(\begin{array}{c} P(y) \wedge \\ g(d(y)) \end{array} \right)$$

(P stratifies x along dimension d with granularity g iff x can be divided into one or more parts in P that are each mapped by d to something in g .)

Here is the theorem along with some relevant corollaries:

(5) **Theorem: $\text{Part}_{\theta,C}$ leads to universal stratified reference**

$$\forall V \forall \theta \forall C \forall C' [C \subseteq C' \rightarrow \text{SR}_{\theta,C'}(\text{Part}_{\theta,C}(V))]$$

(When the Part operator, coindexed with thematic role θ and with granularity threshold C , is applied to any predicate, the result has universal stratified reference with respect to θ and any C' that is at least as coarse as C .)

(6) **Corollary: $\text{Part}_{\theta,C}$ leads to restricted stratified reference**

$$\forall V \forall \theta \forall C \forall C' \forall e [C \subseteq C' \rightarrow \text{SR}_{\theta,C'}(\text{Part}_{\theta,C}(V))(e)]$$

(When the Part operator, coindexed with thematic role θ and with granularity threshold C , is applied to any predicate, the result stratifies any e with respect to θ and any C' that is at least as coarse as C .)

(7) **Corollary: D_{θ} leads to universal stratified reference**

$$\forall V \forall \theta \forall C' [\text{Atom} \subseteq C' \rightarrow \text{SR}_{\theta,C'}(\text{D}_{\theta}(V))]$$

(When the D operator coindexed with thematic role θ is applied to any predicate, the result has universal stratified reference with respect to θ and any C' over an atomic domain.)

(8) **Corollary: D_{θ} leads to restricted stratified reference**

$$\forall V \forall \theta \forall C' \forall e [\text{Atom} \subseteq C' \rightarrow \text{SR}_{\theta,C'}(\text{D}_{\theta}(V))(e)]$$

(When the Part operator coindexed with thematic role θ and with granularity threshold C is applied to any predicate, the result stratifies any e with respect to θ and any C' over an atomic domain.)

To prove Theorem (5), we start with the following tautology:

$$(9) \quad \forall V \forall \theta \forall C \forall e \left[\left[e \in *_{\lambda e'} \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right) \right] \rightarrow \left[e \in *_{\lambda e'} \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right) \right] \right]$$

We rewrite (9) by introducing C' as a superset of C and conjoining $C(\theta(e'))$ with $C'(\theta(e'))$. This is harmless since the first conjunct entails the second:

$$(10) \quad \forall V \forall \theta \forall C \forall C' \forall e \left[\left[C \subseteq C' \wedge e \in *_{\lambda e'} \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right) \right] \right. \\ \left. \rightarrow \left[e \in *_{\lambda e'} \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \wedge C'(\theta(e')) \end{array} \right) \right] \right]$$

We rewrite (10) as follows:

$$(11) \quad \forall V \forall \theta \forall C \forall C'. C \subseteq C' \rightarrow \forall e \left[\left[e \in {}^*\lambda e' \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right) \right] \right. \\ \left. \rightarrow \left[e \in {}^*\lambda e' \left(\left[e' \in \lambda e'' \left(\begin{array}{c} V(e'') \wedge \\ C(\theta(e'')) \end{array} \right) \right] \wedge C'(\theta(e')) \right) \right] \right]$$

From Theorem (19) in Section 2.3.1, we know that $\forall e[V(e) \rightarrow {}^*V(e)]$. Using this fact, we rewrite (11) as follows:

$$(12) \quad \forall V \forall \theta \forall C \forall C'. C \subseteq C' \rightarrow \forall e \left[\left[e \in {}^*\lambda e' \left(\begin{array}{c} V(e') \wedge \\ C(\theta(e')) \end{array} \right) \right] \right. \\ \left. \rightarrow \left[e \in {}^*\lambda e' \left(\left[e' \in {}^*\lambda e'' \left(\begin{array}{c} V(e'') \wedge \\ C(\theta(e'')) \end{array} \right) \right] \wedge C'(\theta(e')) \right) \right] \right]$$

By two applications of the definition of $\text{Part}_{\theta,C}$, we rewrite (12) as follows:

$$(13) \quad \forall V \forall \theta \forall C \forall C'. C \subseteq C' \rightarrow \\ \forall e \left[\text{Part}_{\theta,C}(V)(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{c} \text{Part}_{\theta,C}(V)(e') \wedge \\ C'(\theta(e')) \end{array} \right) \right]$$

Theorem (5) follows from (13) by the definition of universal stratified reference. The corollaries in (6), (7), and (8) follow from it immediately: Restricted stratified reference is a special case of universal stratified reference and is equivalent to the Part operator, and the D operator in turn is a special case of the Part operator.

References

- Abusch, Dorit & Mats Rooth. 1990. Temporal adverbs and the English Perfect. *20th Annual Meeting of the North East Linguistic Society* (NELS 20). 5–19.
- Akmajian, Adrian & Adrienne Lehrer. 1976. NP-like quantifiers and the problem of determining the head of an NP. *Linguistic Analysis* 2(4). 395–413.
- Aldridge, Maurice Vincent. 1982. *English quantifiers: a study of quantifying expressions in linguistic science and modern English usage*. Avebury, UK: Avebury.
- Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 15(1). 5–16. <http://dx.doi.org/10.1007/bf00627432>.
- Bale, Alan. 2009. Yet more evidence for the emptiness of plurality. *38th Annual Meeting of the North East Linguistic Society* (NELS 38). 75–88.
- Bale, Alan Clinton, Michaël Gagnon & Hrayr Khanjian. 2011. Cross-linguistic representations of numerals and number marking. *20th Semantics and Linguistic Theory Conference (SALT 20)*. 582–98. <http://dx.doi.org/10.3765/salt.v20i0.2552>.
- Balusu, Rahul. 2005. Distributive reduplication in Telugu. *36th Annual Meeting of the North East Linguistic Society* (NELS 36). 39–53. <http://semanticsarchive.net/Archive/mEoYTZkZ>.
- Balusu, Rahul & Karattuparambil A. Jayaseelan. 2013. Distributive quantification by reduplication in Dravidian. In Kook-Hee Gil, Stephen Harlow and George Tsoulas (eds), *Strategies of quantification*, 60–86. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199692439.003.0004>.
- Barker, Chris. 1992. Group terms in English: representing groups as atoms. *Journal of Semantics* 9(1). 69–93. <http://dx.doi.org/10.1093/jos/9.1.69>.
- Barker, Chris. 2002. Continuations and the nature of quantification. *Natural Language Semantics* 10(3). 211–42. <http://dx.doi.org/10.1023/a:1022183511876>.
- Barner, David & Jesse Snedeker. 2005. Quantity judgments and individuation: evidence that mass nouns count. *Cognition* 97(1). 41–66. <http://dx.doi.org/10.1016/j.cognition.2004.06.009>.
- Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2). 159–219. <http://dx.doi.org/10.1007/bf00350139>.
- Bayer, Samuel L. 1997. *Confessions of a lapsed Neo-Davidsonian: events and arguments in compositional semantics*. New York: Garland. <http://dx.doi.org/10.4324/9780203055106>.
- Bealer, George. 1979. Predication and matter. In Francis Jeffrey Pelletier (ed.), *Mass terms: some philosophical problems*, 279–94. Berlin: Springer. http://dx.doi.org/10.1007/978-1-4020-4110-5_19.
- Beaver, David. 2001. *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI Publications.
- Beavers, John. 2008. Multiple incremental themes and Figure/Path relations. *18th Semantics and Linguistic Theory Conference (SALT 18)*. 90–107. <http://dx.doi.org/10.3765/salt.v18i0.2495>.
- Beavers, John. 2012. Lexical aspect and multiple incremental themes. In Violeta Demonte & Louise McNally (eds), *Telicity, change, and state: a cross-categorical view of event*

- structure*, 23–59. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199693498.003.0002>.
- Beck, Sigrid & Uli Sauerland. 2000. Cumulation is needed: a reply to Winter 2000. *Natural Language Semantics* 8(4). 349–71. <http://dx.doi.org/10.1023/a:1011240827230>.
- Beck, Sigrid & Arnim von Stechow. 2007. Pluractional adverbials. *Journal of Semantics* 24(3). 215–54. <http://dx.doi.org/10.1093/jos/ffm003>.
- Beghelli, Filippo. 1997. The syntax of distributivity and pair-list readings. In Anna Szabolcsi (ed.), *Ways of scope taking*, 349–408. Dordrecht: Kluwer Academic. http://dx.doi.org/10.1007/978-94-011-5814-5_10.
- Beghelli, Filippo & Tim Stowell. 1997. Distributivity and negation: the syntax of *each* and *every*. In Anna Szabolcsi (ed.), *Ways of scope taking*, 71–107. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-011-5814-5_3.
- Bennett, Michael R. 1974. *Some extensions of a Montague fragment of English*. Los Angeles, CA: University of California PhD thesis.
- Bennett, Michael R. & Barbara H. Partee. 1972. *Toward the logic of tense and aspect in English*. Indiana University Linguistics Club. <http://dx.doi.org/10.1002/9780470751305.ch4>.
- van Benthem, Johan. 1983. *The logic of time*. Dordrecht: Reidel.
- Binnick, Robert I. 1969. *Studies in the derivation of predicative structures*. Chicago: University of Chicago PhD thesis.
- Bittner, Maria. 1994. Cross-linguistic semantics. *Linguistics and Philosophy* 17(1). 53–108. <http://dx.doi.org/10.1007/bf00985041>.
- Blaheta, Don. 2003. *Binominal each: evidence for a modified type system*. Providence, RI: Brown University MA thesis. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.7455&rep=rep1&type=pdf>.
- Bobaljik, Jonathan D. 1995. *Morphosyntax: the syntax of verbal inflection*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. <http://www.ai.mit.edu/projects/dm/theses/bobaljik95.pdf>.
- Bobaljik, Jonathan D. 2001. Floating quantifiers: handle with care. In Lisa Cheng & Rynt Sybesma (eds), *The second state-of-the-article book*, 45–68. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110890952.107>.
- Borik, Olga. 2006. *Aspect and reference time*. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199291298.001.0001>.
- Bortolussi, Bernard. 2013. Latin *quisque* as a “floating” quantifier. *Journal of Latin Linguistics* 12(1). 5–26. <http://dx.doi.org/10.1515/joll-2013-0002>.
- Brasoveanu, Adrian. 2013. Modified numerals as post-suppositions. *Journal of Semantics* 30(2). 155–209. <http://dx.doi.org/10.1093/jos/ffs003>.
- Brasoveanu, Adrian & Jakub Dotlačil. 2015. Strategies for scope taking. *Natural Language Semantics* 23(1). 1–19. <http://dx.doi.org/10.1007/s11050-014-9109-1>.
- Brasoveanu, Adrian & Donka F. Farkas. 2011. How indefinites choose their scope. *Linguistics and Philosophy* 34(1). 1–55. <http://dx.doi.org/10.1007/s10988-011-9092-7>.
- Breheny, Richard. 2008. A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics* 25(2). 93–139. <http://dx.doi.org/10.1093/jos/ffm016>.
- Brisson, Christine. 1998. *Distributivity, maximality and floating quantifiers*. New Brunswick, NJ: Rutgers University PhD thesis. <http://ling.rutgers.edu/images/10.1.1.121.4641.pdf>.

- Brisson, Christine. 2003. Plurals, *all*, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26(2). 129–84. <http://dx.doi.org/10.1023/a:1022771705575>.
- Bunt, Harry C. 1979. Ensembles and the formal semantic properties of mass terms. In Francis Jeffry Pelletier (ed.), *Mass terms: some philosophical problems*, 249–77. Berlin: Springer. http://dx.doi.org/10.1007/978-1-4020-4110-5_18.
- Bunt, Harry C. 1985. *Mass terms and model-theoretic semantics*. Cambridge: Cambridge University Press.
- Bunt, Harry C. 2006. Mass expressions. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn. Amsterdam: Elsevier. <http://dx.doi.org/10.1016/b0-08-044854-2/01097-x>.
- Büring, Daniel. 2005. *Binding theory*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/cb09780511802669>.
- Burzio, Luigi. 1986. *Italian syntax: a government-binding approach*. Berlin: Springer. <http://dx.doi.org/10.1007/978-94-009-4522-7>.
- Cable, Seth. 2014. Distributive numerals and distance distributivity in Tlingit (and beyond). *Language* 90(3). 562–606. <http://dx.doi.org/10.1353/lan.2014.0055>.
- Carlson, Gregory N. 1977. *Reference to kinds in English*. Amherst, MA: University of Massachusetts PhD thesis. <http://semanticsarchive.net/Archive/jk3NzRlY>.
- Carlson, Gregory N. 1984. Thematic roles and their role in semantic interpretation. *Linguistics* 22(3). 259–80. <http://dx.doi.org/10.1515/ling.1984.22.3.259>.
- Carlson, Gregory N. 1989. The semantic interpretation of generic sentences. In Gennaro Chierchia, Barbara H. Partee & Raymond Turner (eds), *Property theory, type theory, and semantics*, vol. 2: *Semantic issues*, 167–92. Dordrecht: Kluwer.
- Carlson, Gregory N. 1998. Thematic roles and the individuation of events. In Susan Rothstein (ed.), *Events and grammar*, 35–51. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-011-3969-4_3.
- Carlson, Gregory N. 2006. Generics, habituals and iteratives. In Keith Brown (ed.), *Encyclopedia of language and linguistics*. Amsterdam: Elsevier. <http://dx.doi.org/10.1016/b0-08-044854-2/00256-x>.
- Carlson, Lauri. 1981. Aspect and quantification. In *Tense and aspect*, 31–64. New York: Academic Press.
- Carnie, Andrew. 2006. *Syntax: a generative introduction*. Oxford: Blackwell.
- Carston, Robyn. 1998. Informativeness, relevance and scalar implicature. In Robin Carston & Seiji Uchida (eds), *Relevance theory: applications and implications*, 179–236. Amsterdam: Benjamins. <http://dx.doi.org/10.1075/pbns.37.11car>.
- Cartwright, Helen. 1975. Amounts and measures of amount. *Noûs* 9(2). 143–64. <http://dx.doi.org/10.2307/2214598>.
- Casati, Roberto & Achille C. Varzi. 1999. *Parts and places: the structures of spatial representation*. Cambridge, MA: MIT Press.
- Champollion, Lucas. 2009. A unified account of distributivity, *for*-adverbials, and measure constructions. In Martin Prinzhorn, Viola Schmitt & Sarah Zobel (eds), *Sinn und Bedeutung* 14, 84–100. Vienna: Universität Wien.
- Champollion, Lucas. 2010a. Cumulative readings of *every* do not provide evidence for events and thematic roles. *Logic, Language and Meaning: 17th Amsterdam Colloquium*. 213–22. http://dx.doi.org/10.1007/978-3-642-14287-1_22.

- Champollion, Lucas. 2010b. *Parts of a whole: distributivity as a bridge between aspect and measurement*. Philadelphia: University of Pennsylvania PhD thesis. <http://repository.upenn.edu/edissertations/958/>.
- Champollion, Lucas. 2011a. Quantification and negation in event semantics. *Baltic international yearbook of cognition, logic and communication* 6. 1–23. <http://dx.doi.org/10.4148/biyclc.v6io.1563>.
- Champollion, Lucas. 2011b. Temporal prepositional phrases and implicit variables. MS. University of Tübingen. <http://www.champollion.com/temporal-pps-and-implicit-variables.pdf>.
- Champollion, Lucas. 2012. *Each vs. jewels: a cover-based view on distance distributivity*. *Logic, language and meaning: 18th Amsterdam Colloquium*. 251–60. http://dx.doi.org/10.1007/978-3-642-31482-7_26.
- Champollion, Lucas. 2013. The scope and processing of *for*-adverbials: a reply to Deo and Piñango. *23rd Semantics and Linguistic Theory Conference (SALT 23)*. 432–52. <http://dx.doi.org/10.3765/salt.v23io.2680>.
- Champollion, Lucas. 2014. Reply to Roger Schwarzschild on event semantics. MS. New York University. <http://ling.auf.net/lingbuzz/002165>.
- Champollion, Lucas. 2015a. *Every boy bought two sausages each: distributivity and dependent numerals*. *32nd West Coast Conference on Formal Linguistics (WCCFL 32)*. 103–10. <http://www.lingref.com/cpp/wccfl/32/abstract3161.html>.
- Champollion, Lucas. 2015b. Refining stratified reference. *Theoretical Linguistics* 41(3–4). 223–40. <http://dx.doi.org/10.1515/tl-2015-0015>.
- Champollion, Lucas. 2015c. Stratified reference: the common core of distributivity, aspect, and measurement. *Theoretical Linguistics* 41(3–4). 109–49. <http://dx.doi.org/10.1515/tl-2015-0008>.
- Champollion, Lucas. 2015d. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy* 38(1). 31–66. <http://dx.doi.org/10.1007/s10988-014-9162-8>.
- Champollion, Lucas. 2016a. Covert distributivity in algebraic event semantics. *Semantics and Pragmatics* 9(15). 1–66. <http://dx.doi.org/10.3765/sp.9.15>.
- Champollion, Lucas. 2016b. Linking the collective-distributive opposition and the telic-atelic opposition. *46th Annual Meeting of the North East Linguistic Society (NELS 46)*, vol. 1, 171–84.
- Champollion, Lucas. 2016c. Overt distributivity in algebraic event semantics. *Semantics and Pragmatics* 9(16). 1–65. <http://dx.doi.org/10.3765/sp.9.16>.
- Champollion, Lucas. 2016d. Ten men and women got married today: noun coordination and the intersective theory of conjunction. *Journal of Semantics*. 33(3). 561–622. <http://dx.doi.org/10.1093/jos/ffvoo8>.
- Champollion, Lucas. to appear. Distributivity, collectivity and cumulativity. In Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds), *Wiley's companion to semantics*. Hoboken, NJ: Wiley. <http://ling.auf.net/lingbuzz/002133>.
- Champollion, Lucas & Manfred Krifka. 2016. Mereology. In Maria Aloni & Paul Dekker (eds), *The Cambridge Handbook of Formal Semantics*, 369–88. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139236157.014>.
- Cheng, Chung-Ying. 1973. Comments on Moravcsik's paper. In Jaakko Hintikka, Julius Moravcsik & Patrick Suppes (eds), *Approaches to natural language*, 286–8. Dordrecht: Reidel. http://dx.doi.org/10.1007/978-94-010-2506-5_14.

- Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of “semantic parameter.” In Susan Rothstein (ed.), *Events and grammar*, 53–104. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-011-3969-4_4.
- Chierchia, Gennaro. 1998b. Reference to kinds across language. *Natural Language Semantics* 6(4). 339–405. <http://dx.doi.org/10.1023/a:1008324218506>.
- Chierchia, Gennaro. 2006. Broaden your views: implications of domain widening and the “logicality” of language. *Linguistic Inquiry* 37(4). 535–90. <http://dx.doi.org/10.1162/ling.2006.37.4.535>.
- Chierchia, Gennaro. 2008. *On being “one.”* Paper presented at the 2008 Workshop on Markedness and Underspecification in the Morphology and Semantics of Agreement (MUMSA).
- Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174(1). 99–149. <http://dx.doi.org/10.1007/s11229-009-9686-6>.
- Choe, Jae-Woong. 1987. *Anti-quantifiers and a theory of distributivity*. Amherst, MA: University of Massachusetts PhD thesis. <http://scholarworks.umass.edu/dissertations/AAI8727027/>.
- Choe, Jae-Woong. 1991. Georgian reduplication and a relational theory of distributivity. In Akira Ikeya (ed.), *Sixth Japanese–Korean Joint Conference on Formal Linguistics*. Tokyo: Logico-Linguistic Society of Japan. <https://dspace.wul.waseda.ac.jp/dspace/bitstream/2065/11793/1/JK6-4-25.pdf>.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris. <http://dx.doi.org/10.1515/9783110884166>.
- Chomsky, Noam. 1982. *The generative enterprise. A discussion with Riny Huybregts and Henk van Riemsdijk*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics* 20. 104–131.
- Cohen, Ariel. 2005. More than bare existence: an implicature of existential bare plurals. *Journal of Semantics* 22. 389–400. <http://dx.doi.org/10.1093/jos/ffh031>.
- Collins, Chris & Paul Martin Postal. 2012. *Imposters: a study of pronominal agreement*. Cambridge, MA: MIT Press. <http://dx.doi.org/10.7551/mitpress/9780262016889.001.0001>.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Corver, Norbert. 2015. (Un)boundedness across syntactic categories. *Theoretical Linguistics* 41(3–4). 151–165. <http://dx.doi.org/10.1515/tl-2015-0009>.
- Crain, Stephen. 2011. Meaning in first language acquisition. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds), *Semantics: an international handbook of natural language meaning*, 2724–52. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110253382.2724>.
- Cresswell, Max J. 1976. The semantics of degree. In Barbara H. Partee (ed.), *Montague grammar*, 261–92. New York: Academic Press. <http://www.sfs.uni-tuebingen.de/~astechow/Lehre/Pariso8/cresswell76%20Kopie.pdf>
- Cresswell, Max J. 1985. Review of Landman and Veltman (eds), 1984, *Varieties of Formal Semantics*. *Linguistics* 23.
- Daubert, B. F. & T. H. Clarke. 1944. Unsaturated synthetic glycerides. VI. Polymorphism of symmetrical monoöleyldisaturated triglycerides. *Journal of the American Chemical Society* 66(5). 690–91. <http://dx.doi.org/10.1021/jao1233a008>.

- Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The logic of decision and action*, 81–95. Pittsburgh, PA: University of Pittsburgh Press. <http://dx.doi.org/10.1093/0199246270.003.0006>.
- Davidson, Donald. 1969. The individuation of events. In Nicholas Rescher (ed.), *Essays in honor of Carl G. Hempel: a tribute on the occasion of his sixty-fifth birthday*, 216–34. Dordrecht: Springer. http://dx.doi.org/10.1007/978-94-017-1466-2_11.
- Dekker, Paul. 1997. Cases, adverbs, situations and events. In Hans Kamp & Barbara H. Partee (eds), *Workshop on context dependence*. IMS, Stuttgart & UFAL, Prague. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.16.7314&rep=rep1&type=pdf>.
- Delfitto, Denis. 2005. Bare plurals. In Martin Everaert & Henk van Riemsdijk (eds), *The Blackwell companion to syntax*. Oxford: Blackwell. <http://dx.doi.org/10.1002/9780470996591.ch8>.
- Deo, Ashwini Sharad. 2009. Unifying the imperfective and the progressive: partitions as quantificational domains. *Linguistics and Philosophy* 32(5). 475–521. <http://dx.doi.org/10.1007/s10988-010-9068-z>.
- Deo, Ashwini Sharad & Maria Mercedes Piñango. 2011. Quantification and context in measure adverbs. *21st Semantics and Linguistic Theory Conference (SALT 21)*. 295–312. <http://dx.doi.org/10.3765/salt.v21i0.2614>.
- Dobrovie-Sorin, Carmen. 2014. Collective quantification and the homogeneity constraint. *24th Semantics and Linguistic Theory Conference (SALT 24)*. 453–72. <http://dx.doi.org/10.3765/salt.v24i0.2428>.
- van der Does, Jaap. 1993. Sums and quantifiers. *Linguistics and Philosophy* 16(5). 509–50. http://dx.doi.org/10.1007/978-94-017-2706-8_6.
- van der Does, Jaap & Henk J. Verkuyl. 1996. Quantification and predication. In Kees van Deemter & Stanley Peters (eds), *Semantic ambiguity and underspecification*, 27–54. Stanford, CA: CSLI Publications. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.378.829&rep=rep1&type=pdf>.
- Doetjes, Jenny. 1997. *Quantifiers and selection: on the distribution of quantifying expressions in French, Dutch and English*. The Hague: Holland Institute of Generative Linguistics PhD thesis. https://openaccess.leidenuniv.nl/bitstream/handle/1887/19731/doetjes_geheel%5B1%5D.pdf?sequence=1.
- Doetjes, Jenny. 2015. Very small or just not too big? Stratified reference and granularity. *Theoretical Linguistics* 41(3–4). 167–76. <http://dx.doi.org/10.1515/tl-2015-0010>.
- Dotlačil, Jakub. 2011. Fastidious distributivity. *21st Semantics and Linguistic Theory Conference (SALT 21)*. 313–32. <http://dx.doi.org/10.3765/salt.v21i0.2596>.
- Dotlačil, Jakub. 2012. Binominal *each* as an anaphoric determiner: compositional analysis. In Ana Aguilar Guevara, Anna Chernilovskaya & Rick Nouwen (eds), *Sinn und Bedeutung* 16, 211–24. Cambridge: MIT Working Papers in Linguistics. <http://mitwpl.mit.edu/open/sub16/Dotlacil.pdf>.
- Dotlačil, Jakub. 2013. Reciprocals distribute over information states. *Journal of Semantics* 30(4). 423–77. <http://dx.doi.org/10.1093/jos/ffso16>.
- Dougherty, Ray C. 1970. A grammar of coordinate conjoined structures: I. *Language* 46(4). 850–98. <http://dx.doi.org/10.2307/412260>.
- Dougherty, Ray C. 1971. A grammar of coordinate conjoined structures: II. *Language* 47(2). 298–339. <http://dx.doi.org/10.2307/412083>.

- Dowty, David R. 1972. *Studies in the logic of verb aspect and time reference in English*. Austin: University of Texas PhD thesis.
- Dowty, David R. 1979. *Word meaning and Montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: Reidel. <http://dx.doi.org/10.1007/978-94-009-9473-7>.
- Dowty, David R. 1987. Collective predicates, distributive predicates, and *all*. In Ann Miller & Zheng-sheng Zhang (eds), *3rd Eastern States Conference on Linguistics (ESCOL)*, 97–115. Columbus: Ohio State University.
- Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619. <http://dx.doi.org/10.1353/lan.1991.0021>.
- Dowty, David R. & Belinda Brodie. 1984. The semantics of “floated” quantifiers in a transformationless grammar. *3rd West Coast Conference on Formal Linguistics (WCCFL)* 3. 75–90.
- Eberle, Kurt. 1998. The influence of Plural NPs on the Aktionsart in DRT. In Fritz Hamm & Erhard Hinrichs (eds), *Plural quantification*. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-017-2706-8_3.
- Eckardt, Regine. 1998. *Adverbs, events, and other things*. Tübingen: Niemeyer. <http://dx.doi.org/10.1515/9783110913781>.
- Egg, Markus. 1994. *Aktionsart und Kompositionalität*. Berlin: Akademie.
- Egg, Markus. 1995. Aspect and quantification: an iterative approach. *10th Amsterdam Colloquium*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.37.2666&rep=rep1&type=pdf>.
- Erlewine, Michael Yoshitaka & Hadas Kotek. 2016. A streamlined approach to online linguistic surveys. *Natural Language and Linguistic Theory* 34(2). 481–95. <http://dx.doi.org/10.1007/s11049-015-9305-9>. <http://dx.doi.org/10.1007/s11049-015-9305-9>.
- Farkas, Donka F. 1997. Dependent indefinites. In Francis Corblin, Daniele Godard & Jean-Marie Marandin (eds), *Empirical issues in formal syntax and semantics*, 243–68. Bern: Peter Lang. <http://people.ucsc.edu/~farkas/papers/reduplication.pdf>.
- Farkas, Donka F. & Henriette de Swart. 2010. The semantics and pragmatics of plurals. *Semantics and Pragmatics* 3(6). 1–54. <http://dx.doi.org/10.3765/sp.3.6>.
- Ferreira, Marcelo. 2005. *Event quantification and plurality*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. <http://hdl.handle.net/1721.1/33697>.
- Filip, Hana. 1999. *Aspect, eventuality types, and nominal reference*. New York: Garland.
- Filip, Hana. 2000. The quantization puzzle. In James Pustejovsky & Carol Tenny (eds), *Events as grammatical objects, from the combined perspectives of lexical semantics, logical semantics and syntax*, 3.60. Stanford, CA: CSLI Publications.
- Fine, Kit. 2012. Counterfactuals without possible worlds. *Journal of Philosophy* 109(3). 221–46. <http://dx.doi.org/10.5840/jphil201210938>.
- Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–86. <http://dx.doi.org/10.1007/s10988-006-9004-4>.
- Gärdenfors, Peter. 2007. Representing actions and functional properties in conceptual spaces. In Tom Ziemke, Jordan Zlatev & Roslyn M. Frank (eds), *Body, language and mind*, volume 1: *Embodiment*, 167–96. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110207507.2.167>.
- Garey, Howard B. 1957. Verbal aspect in French. *Language* 33(2). 91–110. <http://dx.doi.org/10.2307/410722>.

- Gawron, Jean Mark. 2002. *Two kinds of determiners in DP*. Talk at LSA winter meeting. http://gawron.sdsu.edu/lsa_hand.ps.
- Gawron, Jean Mark. 2004. Motion, scalar paths, and lexical aspect. Paper presented at Workshop on Event Structures in Linguistic Form and Interpretation, Leipzig. http://gawron.sdsu.edu/new_motion_to_state_handout.pdf.
- Gawron, Jean Mark. 2005. Generalized paths. *15th Semantics and Linguistic Theory Conference (SALT 15)*. http://gawron.sdsu.edu/salt_paper.pdf.
- Gawron, Jean Mark. 2009. *The lexical semantics of extent verbs*. Presented at San Diego State University, January 31.
- van Geenhoven, Veerle. 2004. For-adverbials, frequentative aspect, and pluractionality. *Natural Language Semantics* 12(2). 135–90. <http://dx.doi.org/10.1023/b:nals.0000031389.36427.af>.
- van Geenhoven, Veerle. 2005. Atelicity, pluractionality and adverbial quantification. In Henk J. Verkuyl, Henriette de Swart & Angeliek van Hout (eds), *Perspectives on aspect*, 107–24. Berlin: Springer. http://dx.doi.org/10.1007/1-4020-3232-3_6.
- Geurts, Bart. 2006. Take “five”: the meaning and use of a number word. In Svetlana Vogelev & Liliane Tasmowski (eds), *Non-definiteness and plurality*, 311–29. Amsterdam: Benjamins. <http://dx.doi.org/10.1075/la.95.16geu>.
- Gil, David. 1982. *Distributive numerals*. Los Angeles: University of California PhD thesis.
- Gil, David. 1989. Nouns, verbs, and quantification. MS. Tel Aviv University.
- Gil, David. 1993. Nominal and verbal quantification. *STUF: Language Typology and Universals* 46(1–4). 275–317. <http://dx.doi.org/10.1524/stuf.1993.46.14.275>.
- Gil, David. 1995. Universal quantifiers and distributivity. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara Hall Partee (eds), *Quantification in natural languages*, 321–62. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-017-2817-1_11.
- Gil, David. 2013. Distributive numerals. In Matthew S. Dryer & Martin Haspelmath (eds), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/54>.
- Gillon, Brendan S. 1987. The readings of plural noun phrases in English. *Linguistics and Philosophy* 10(2). 199–219. <http://dx.doi.org/10.1007/bfo0584318>.
- Gillon, Brendan S. 1990. Plural noun phrases and their readings: a reply to Lasnik. *Linguistics and Philosophy* 13(4). 477–85. <http://dx.doi.org/10.1007/bfo0630751>.
- Gillon, Brendan S. 1992. Towards a common semantics for English mass and count nouns. *Linguistics and Philosophy* 15(6). 597–640. <http://dx.doi.org/10.1007/bfo0628112>.
- Graham, Daniel W. 1980. States and performances: Aristotle’s test. *Philosophical Quarterly* 30(119). 117–30. <http://dx.doi.org/10.2307/2219276>.
- Grevisse, Maurice. 1980. *Le bon usage: grammaire française, avec des remarques sur la langue française d’aujourd’hui*. Paris: Duculot.
- Groenendijk, Jeroen, Theo Janssen & Martin Stokhof (eds). 1984. *Truth, interpretation and information: 3rd Amsterdam Colloquium*. Dordrecht: Foris. <http://dx.doi.org/10.1515/9783110867602>.
- Gruber, Jeffrey S. 1965. *Studies in lexical relations*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. <http://hdl.handle.net/1721.1/13010>.
- Guéron, Jacqueline. 1979. On the structure of “pseudo-partitive” NP’s. *Bulletin of the ICU Summer Institute in Linguistics* 12. 21–48.

- Hackl, Martin. 2001. *Comparative quantifiers*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. <http://hdl.handle.net/1721.1/8765>.
- Hackl, Martin. 2002. The ingredients of essentially plural predicates. *32nd Annual Meeting of the North East Linguistic Society (NELS 32)*. 171–82. <http://web.mit.edu/hackl/www/papers/files/HacklNelson1.pdf>.
- Hackl, Martin. 2009. On the grammar and processing of proportional quantifiers: most versus more than half. *Natural Language Semantics* 17(1). 63–98. <http://dx.doi.org/10.1007/s11050-008-9039-x>.
- Harnish, Robert M. 1976. Logical form and implicature. In Thomas G. Bever, Jerrold J. Katz & D. Terence Langendoen (eds), *An integrated theory of linguistic ability*, 313–92. New York: Crowell.
- Hausser, Roland. 1974. *Quantification in an extended Montague grammar*. Austin: University of Texas PhD thesis.
- Heim, Irene. 1994. Plurals. MS. Massachusetts Institute of Technology. http://www.sfs.uni-tuebingen.de/~astechow/Lehre/SS08/Heim_Plurals%20'94.pdf.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Heim, Irene, Howard Lasnik & Robert May. 1991. Reciprocity and plurality. *Linguistic Inquiry* 22(1). 63–101. <http://www.jstor.org/stable/4178708>.
- Hellan, Lars. 1981. *Towards an integrated analysis of comparatives*. Tübingen: Narr.
- Hendriks, Herman. 1993. *Studied flexibility*. Amsterdam: University of Amsterdam PhD thesis.
- Heny, Frank. 1982. Tense, aspect and time adverbials, part II. *Linguistics and Philosophy* 5(1). 109–54. <http://dx.doi.org/10.1007/bf00390694>.
- Higginbotham, James. 1994. Mass and count quantifiers. *Linguistics and Philosophy* 17(5). 447–80. <http://dx.doi.org/10.1007/bf00985831>.
- Hinrichs, Erhard. 1985. *A compositional semantics for aktionsarten and NP reference in English*. Columbus: Ohio State University PhD thesis. https://etd.ohiolink.edu/!etd.send_file?accession=osu1272461401&disposition=inline.
- Hitzeman, Janet. 1991. Aspect and adverbials. *1st Semantics and Linguistic Theory Conference (SALT 1)*. 107–27. <http://dx.doi.org/10.3765/salt.v1i0.2763>.
- Hitzeman, Janet. 1997. Semantic partition and the ambiguity of sentences containing temporal adverbials. *Natural Language Semantics* 5(2). 87–100. <http://dx.doi.org/10.1023/a:1008221528793>.
- Hoeksema, Jack. 1983. Plurality and conjunction. In Alice ter Meulen (ed.), *Studies in model-theoretic semantics*, 63–83. Dordrecht: Foris.
- Hogg, Richard M. 1977. *English quantifier systems*. Amsterdam: North-Holland.
- Horn, Laurence. 1972. *On the semantics and pragmatics of logical operators in English*. Los Angeles: University of California PhD thesis.
- Horn, Laurence. 1989. *A natural history of negation*. Stanford, CA: CSLI Publications.
- Hovda, Peter. 2009. What is classical mereology? *Journal of Philosophical Logic* 38(1). 55–82. <http://dx.doi.org/10.1007/s10992-008-9092-4>.
- Iatridou, Sabine, Elena Anagnostopoulou & Roumyana Izvorski. 2001. Observations about the form and meaning of the Perfect. In *Ken Hale: a life in language*, 189–238. Cambridge, MA: MIT Press. <http://dx.doi.org/10.1515/9783110902358.153>.

- Ionin, Tania & Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23(4). 315–60. <http://dx.doi.org/10.1093/jos/ffl006>.
- Ionin, Tania, Ora Matushansky & Eddy G. Ruys. 2006. Parts of speech: Towards a unified semantics for partitives. *36th Annual Meeting of the North East Linguistic Society (NELS 36)*. 357–70. http://www.researchgate.net/publication/46682750_Parts_of_Speech_Toward_a_Unified_Semantics_for_Partitives.
- Iordăchioaia, Gianina & Elena Soare. 2015. Pluractionality with lexically cumulative verbs. *Natural Language Semantics* 23(4). 307–52. <http://dx.doi.org/10.1007/s11050-015-9117-9>.
- Ioup, Georgette. 1975. Some universals for quantifier scope. *Syntax and Semantics* 4. 37–58.
- Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jackendoff, Ray S. 1996. The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory* 14(2). 305–54. <http://dx.doi.org/10.1007/bf00133686>.
- Janke, Vikki & Ad Neeleman. 2012. Ascending and descending VPs in English. *Linguistic Inquiry* 43(1). 151–90. http://dx.doi.org/10.1162/ling_a_00082.
- Junker, Marie-Odile. 1995. *Syntaxe et sémantique des quantifieurs flottants tous et chacun: distributivité en sémantique conceptuelle*. Geneva: Droz.
- Kamp, Hans. 1979. Events, instants and temporal reference. In Rainer Bäuerle, Urs Egli & Arnim von Stechow (eds), *Semantics from different points of view*, 376–417. Berlin: Springer. http://dx.doi.org/10.1007/978-3-642-67458-7_24.
- Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht, Netherlands: Kluwer. <http://dx.doi.org/10.1007/978-94-017-1616-1>.
- Karttunen, Lauri. 1974. Until. In *Papers from the 10th regional meeting of the Chicago Linguistic Society (CLS 10)*, 283–97. Chicago. <http://web.stanford.edu/~laurik/publications/archive/until.pdf>.
- Karttunen, Lauri & Stanley Peters. 1979. Conventional implicature. In Choon-Kyu Oh & David A. Dinneen (eds), *Presupposition*, 1–56. New York: Academic Press.
- Keim, Marybelle C. & John P. Murray. 2008. Chief academic officers' demographics and educational backgrounds. *Community College Review* 36(2). 116–32. <http://dx.doi.org/10.1177/0091552108324657>.
- Keizer, Evelien. 2007. *The English noun phrase: the nature of linguistic categorization*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/cbo9780511627699>.
- Kennedy, Christopher. 2007. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1). 1–45. <http://dx.doi.org/10.1007/s10988-006-9008-0>.
- Kennedy, Christopher. 2009. The number of meanings of English number words. MS. University of Chicago. <http://semantics.uchicago.edu/kennedy/docs/uiuc-slides.pdf>.
- Kiparsky, Paul. 1998. Partitive case and aspect. In Miriam Butt & Wilhelm Geuder (eds), *The projection of arguments: lexical and compositional factors*, 265–307. Stanford, CA: CSLI Publications. <http://web.stanford.edu/~kiparsky/Papers/wuppertal.pdf>.
- Kipper-Schuler, Karin. 2005. *VerbNet: a broad-coverage, comprehensive verb lexicon*. Philadelphia: University of Pennsylvania PhD thesis. <http://verbs.colorado.edu/~kipper/Papers/dissertation.pdf>.

- Kleiber, Georges. 1989. «Le» générique: un massif? *Langages* 94. 73–113. <http://www.jstor.org/stable/41682171>.
- Klein, Ewan. 1991. Comparatives. In Dieter Wunderlich & Arnim von Stechow (eds), *Semantik: ein internationales Handbuch der zeitgenössischen Forschung / Semantics: an international handbook of contemporary research*, 673–91. Berlin: de Gruyter.
- Krantz, David H., R. Duncan Luce, Patrick Suppes & Amos Tversky. 1971. *Foundations of measurement*. New York: Academic Press.
- Krasikova, Sveta. 2010. Scales in the meaning of adjectives. In A. Grønn & I. Marijanovic (eds), *Russian in contrast*, 51–74. Oslo: Oslo Studies in Language.
- Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12(5). 607–53. <http://dx.doi.org/10.1007/bf00627775>.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In J. Rooryck & L. Zaring (eds), *Phrase structure and the lexicon*, 109–37. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-015-8617-7_5.
- Kratzer, Angelika. 2000. The event argument and the semantics of verbs (ch. 2). MS. University of Massachusetts, Amherst. <http://semanticsarchive.net/Archive/GU1NWM4Z/>.
- Kratzer, Angelika. 2001. The event argument and the semantics of verbs (ch. 4). MS. University of Massachusetts, Amherst. <http://semanticsarchive.net/Archive/GU1NWM4Z/>.
- Kratzer, Angelika. 2003. The event argument and the semantics of verbs (ch. 3). MS. University of Massachusetts, Amherst. <http://semanticsarchive.net/Archive/GU1NWM4Z/>.
- Kratzer, Angelika. 2007. On the plurality of verbs. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds), *Event structures in linguistic form and interpretation*, 269–300. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110925449.269>.
- Kratzer, Angelika. 2016. Situations in natural language semantics. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2016. Stanford, CA. <http://plato.stanford.edu/archives/spr2016/entries/situations-semantics/>.
- Krifka, Manfred. 1989a. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds), *Semantics and contextual expression*, 75–115. Dordrecht: Foris. <http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka%201989%20Nominal%20Reference.pdf>.
- Krifka, Manfred. 1989b. *Nominalreferenz und Zeitkonstitution: zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. Munich: Fink. http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka_1989_Nominalreferenz_Zeitkonstitution.PDF.
- Krifka, Manfred. 1990. Four thousand ships passed through the lock: object induced measure functions on events. *Linguistics and Philosophy* 13(5). 487–520. <http://dx.doi.org/10.1007/bf00627291>.
- Krifka, Manfred. 1991. Massennomina. In Arnim von Stechow & Dieter Wunderlich (eds), *Semantik: ein internationales Handbuch der zeitgenössischen Forschung / Semantics: an international handbook of contemporary research*, 399–417. Berlin: de Gruyter.
- Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Ivan A. Sag & Anna Szabolcsi (eds), *Lexical matters*, 29–53. Stanford, CA: CSLI Publications. <http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka%201992%20Thematic%20Relations.PDF>.
- Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (ed.), *Events and grammar*, 197–235. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-011-3969-4_9.

- Krifka, Manfred. 1999. At least some determiners aren't determiners. In Ken Turner (ed.), *The semantics/pragmatics interface from different points of view*, 257–91. Amsterdam: Elsevier. <http://amor.cms.hu-berlin.de/h2816i3x/Publications/ATLEASTSOME.pdf>.
- Krifka, Manfred. 2004. Bare NPs: kind-referring, indefinites, both, or neither? *13th Semantics and Linguistic Theory Conference (SALT 13)*. 180–203. <http://dx.doi.org/10.3765/salt.v13i0.2880>.
- Krifka, Manfred, Francis Jeffry Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia & Godehard Link. 1995. Genericity: an introduction. In Gregory N. Carlson & Francis Jeffry Pelletier (eds), *The generic book*, 1–124. Chicago: University of Chicago Press.
- Križ, Manuel. 2016. Homogeneity, non-maximality, and *all*. *Journal of Semantics* 33(3): 493–538. <http://dx.doi.org/10.1093/jos/ffv006>.
- Kroch, Anthony S. 1974. *The semantics of scope in English*. Cambridge: Massachusetts Institute of Technology PhD thesis. <http://hdl.handle.net/1721.1/13020>.
- Kuhn, Jeremy. 2014. *Gather-type predicates: massiness over participants*. Presentation at the 45th Annual Meeting of the North East Linguistic Society (NELS 45). <http://www.jeremykuhn.net/papers/Kuhn-gather-slides.pdf>.
- Kuhn, Jeremy. 2015. *Cross-categorical singular and plural reference in sign language*. New York: University PhD thesis. <http://www.jeremykuhn.net/papers/Kuhn-dissertation.pdf>.
- Kuno, Susumu. 1982. The focus of the question and the focus of the answer. In R. Schneider, K. Tuite, and R. Chametzky (eds), *Papers from the parasession on nondeclaratives*, 134–57. Chicago: Chicago Linguistics Society.
- Ladusaw, William. 1982. Semantic constraints on the English partitive construction. *1st West Coast Conference on Formal Linguistics (WCCFL 1)*. 231–42.
- Landman, Fred. 1989. Groups, I. *Linguistics and Philosophy* 12(5). 559–605. <http://dx.doi.org/10.1007/bf00627774>.
- Landman, Fred. 1991. *Structures for semantics*. Dordrecht: Kluwer.
- Landman, Fred. 1996. Plurality. In Shalom Lappin (ed.), *Handbook of contemporary semantic theory*, 425–57. Oxford: Blackwell. <http://dx.doi.org/10.1111/b.9780631207498.1997.00019.x>.
- Landman, Fred. 2000. *Events and plurality: the Jerusalem lectures*. Dordrecht: Kluwer. <http://dx.doi.org/10.1007/978-94-011-4359-2>.
- Landman, Fred. 2004. *Indefinites and the type of sets*. Oxford: Blackwell. <http://dx.doi.org/10.1002/9780470759318>.
- Landman, Fred & Susan Rothstein. 2009. Incremental homogeneity in the semantics of aspectual *for*-phrases. In Malka Rappaport Hovav, Ivi Sichel & Edit Doron (eds), *Syntax, lexical semantics and event structure*, 229–51. Amsterdam: Benjamins. <http://dx.doi.org/10.1093/acprof:oso/9780199544325.003.0011>.
- Landman, Fred & Susan Rothstein. 2012a. The felicity of aspectual *for*-phrases, part 1: Homogeneity. *Language and Linguistics Compass* 6(2). 85–96. <http://dx.doi.org/10.1002/ln3.324>.
- Landman, Fred & Susan Rothstein. 2012b. The felicity of aspectual *for*-phrases, part 2: Incremental homogeneity. *Language and Linguistics Compass* 6(2). 97–112. <http://dx.doi.org/10.1002/ln3.323>.
- Langacker, Ronald. 1986. *Foundations of cognitive grammar*. Stanford, CA: Stanford University Press.
- Lasersohn, Peter. 1988. *A semantics for groups and events*. Columbus: Ohio State University PhD thesis. <http://semanticsarchive.net/Archive/zQ2ODQxN/>.

- Lasnik, Peter. 1989. On the readings of plural noun phrases. *Linguistic Inquiry* 20(1). 130–4. <http://www.jstor.org/stable/4178619>.
- Lasnik, Peter. 1992. Generalized conjunction and temporal modification. *Linguistics and Philosophy* 15(4). 381–410. <http://dx.doi.org/10.1007/bf00627682>.
- Lasnik, Peter. 1995. *Plurality, conjunction and events*. Dordrecht: Kluwer. <http://dx.doi.org/10.1007/978-94-015-8581-1>.
- Lasnik, Peter. 1998a. Events in the semantics of collectivizing adverbials. In Susan Rothstein (ed.), *Events and grammar*, 273–92. Dordrecht: Springer. http://dx.doi.org/10.1007/978-94-011-3969-4_11.
- Lasnik, Peter. 1998b. Generalized distributivity operators. *Linguistics and Philosophy* 21(1). 83–93. <http://dx.doi.org/10.1023/a:1005317815339>.
- Lasnik, Peter. 1999. Pragmatic halos. *Language* 75(3). 522–51. <http://dx.doi.org/10.2307/417059>.
- Lasnik, Peter. 2011. Mass nouns and plurals. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds), *Semantics: an international handbook of natural language meaning*, 1131–53. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110255072.1131>.
- Lassiter, Daniel. 2010a. Explaining a restriction on the scope of the comparative operator. *University of Pennsylvania Working Papers in Linguistics* 16(1). <http://repository.upenn.edu/cgi/viewcontent.cgi?article=1107&context=wppl>.
- Lassiter, Daniel. 2010b. The algebraic structure of amounts: evidence from comparatives. In Thomas Icard & Reinhard Muskens (eds), *Interfaces: explorations in logic, language, and computation (LNCS 6211)*. Berlin: Springer. http://dx.doi.org/10.1007/978-3-642-14729-6_4.
- LaTerza, Chris. 2014a. *Distributivity and plural anaphora*. College Park: University of Maryland PhD thesis. <http://hdl.handle.net/1903/15826>.
- LaTerza, Chris. 2014b. Local plural anaphora as sub-event distributivity. *32nd West Coast Conference on Formal Linguistics (WCCFL 32)*. 141–8. <http://www.lingref.com/cpp/wccfl/32/paper3165.pdf>.
- Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- Levinson, Stephen C. 2000. *Presumptive meanings*. Cambridge, MA: MIT Press.
- Lewis, David. 1972. General semantics. In Donald Davidson & Gilbert Harman (eds), *Semantics of natural language*, 169–218. Dordrecht: Reidel. http://dx.doi.org/10.1007/978-94-010-2557-7_7.
- Lin, Jo-Wang. 1998. Distributivity in Chinese and its implications. *Natural Language Semantics* 6(2). 201–43. <http://dx.doi.org/10.1023/a:1008299031574>.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Reiner Bäuerle, Christoph Schwarze & Arnim von Stechow (eds), *Meaning, use and interpretation of language*, 303–23. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110852820.302>. Repr. in Link, *Algebraic semantics in language and philosophy*, ch. 1 (Stanford, CA: CSLI Publications, 1998).
- Link, Godehard. 1987a. Algebraic semantics of event structures. 6th Amsterdam Colloquium. Repr. in Link, *Algebraic semantics in language and philosophy*, ch. 11 (Stanford, CA: CSLI Publications, 1998).
- Link, Godehard. 1987b. Generalized quantifiers and plurals. In Peter Gärdenfors (ed.), *Generalized quantifiers*, 151–80. Berlin: Springer. http://dx.doi.org/10.1007/978-94-009-3381-1_6.

- Repr. in Link, *Algebraic semantics in language and philosophy*, ch. 4 (Stanford, CA: CSLI Publications, 1998).
- Link, Godehard. 1991a. Plural. In Arnim von Stechow & Dieter Wunderlich (eds), *Semantik: ein internationales Handbuch der zeitgenössischen Forschung / Semantics: An international handbook of contemporary research*, 418–40. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110126969.6.418>. Repr. in Link, *Algebraic semantics in language and philosophy*, ch. 2 (Stanford, CA: CSLI Publications, 1998).
- Link, Godehard. 1991b. Quantity and number. In Dietmar Zaefferer (ed.), *Semantic universals and universal semantics*, 133–49. Dordrecht: Foris. Repr. in Link, *Algebraic semantics in language and philosophy*, ch. 9 (Stanford, CA: CSLI Publications, 1998).
- Link, Godehard. 1997. Ten years of research on plurals—where do we stand? In Fritz Hamm & Erhard Hinrichs (eds), *Plurality and quantification*, 19–54. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-017-2706-8_2. Repr. in Link, *Algebraic semantics in language and philosophy*, ch. 7 (Stanford, CA: CSLI Publications, 1998).
- Link, Godehard. 1998a. *Algebraic semantics in language and philosophy*. Stanford, CA: CSLI Publications.
- Link, Godehard. 1998b. *Je drei Äpfel—three apples each*: quantification and the German *je*. In Link 1998, *Algebraic semantics in language and philosophy*, ch. 5, 117–32. Stanford, CA: CSLI Publications.
- Link, Godehard. 2015. Stratified reference: a note on its logic. *Theoretical Linguistics* 41(3–4). 177–83. <http://dx.doi.org/10.1515/tl-2015-0011>.
- Lønning, Jan Tore. 1987. Mass terms and quantification. *Linguistics and Philosophy* 10(1). 1–52. <http://dx.doi.org/10.1007/bf00603391>.
- Malamud, Sophia A. 2006a. (Non)maximality and distributivity: a decision theory approach. *16th Semantics and Linguistic Theory Conference (SALT 16)*. 120–37. <http://dx.doi.org/10.3765/salt.v16i0.2943>.
- Malamud, Sophia A. 2006b. *Semantics and pragmatics of arbitrariness*. Philadelphia: University of Pennsylvania PhD thesis. <http://people.brandeis.edu/~smalamud/Malamud-diss-single.pdf>.
- Malamud, Sophia A. 2012. The meaning of plural definites: a decision-theoretic approach. *Semantics and Pragmatics* 5(3). 1–58. <http://dx.doi.org/10.3765/sp.5.3>.
- Maling, Joan M. 1976. Notes on quantifier-postposing. *Linguistic Inquiry* 7(4). 708–18. <http://www.jstor.org/stable/4177962>.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge: MIT Press. <http://www.ai.mit.edu/projects/dm/theses/marantz81.pdf>.
- Markosian, Ned. 2014. Time. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2014. Stanford, CA. <http://plato.stanford.edu/archives/spr2014/entries/time/>.
- Matthewson, Lisa. 2001. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* 9(2). 145–89. <http://dx.doi.org/10.1023/a:1012492911285>.
- McCawley, James D. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4(3). 243–69. <http://www.jstor.org/stable/25000330>.
- de Mey, Sjaap. 1981. The dependent plural and the analysis of tense. *11th Annual Meeting of the North East Linguistic Society (NELS 11)*.
- Mills, Ian, Tomislav Cvitaš, Klaus Homann, Nikola Kallay & Kozo Kuchitsu (eds). 2007. *Quantities, units and symbols in physical chemistry (IUPAC Green Book)*. Oxford: Blackwell Science.

- Milsark, Gary. 1974. *Existential sentences in English*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. <http://dspace.mit.edu/handle/1721.1/13021>.
- Mittwoch, Anita. 1982. On the difference between *eating* and *eating something*: activities versus accomplishments. *Linguistic Inquiry* 13(1). 113–22. <http://www.jstor.org/stable/4178263>.
- Mittwoch, Anita. 1988. Aspects of English aspect: on the interaction of perfect, progressive and durational phrases. *Linguistics and Philosophy* 11(2). 203–54. <http://dx.doi.org/10.1007/bfoo632461>.
- Mollá-Aliod, Diego. 1997. *Aspectual composition and sentence interpretation: a formal approach*. Edinburgh: University of Edinburgh PhD thesis. <http://comp.mq.edu.au/~diego/publications/phd/thesis.pdf>.
- Moltmann, Friederike. 1989. Measure adverbials as part quantifiers. *8th West Coast Conference on Formal Linguistics (WCCFL 8)*. 247–61. <http://semantics.univ-parisi.fr/pdf/Measure%20Adverbials.pdf>.
- Moltmann, Friederike. 1991. Measure adverbials. *Linguistics and Philosophy* 14(6). 629–60. <http://dx.doi.org/10.1007/bfoo631962>.
- Moltmann, Friederike. 1992. Reciprocals and *same/different*: towards a semantic analysis. *Linguistics and Philosophy* 15(4). 411–62. <http://dx.doi.org/10.1007/bfoo627683>.
- Moltmann, Friederike. 1997. *Parts and wholes in semantics*. Oxford: Oxford University Press.
- Moltmann, Friederike. 1998. Part structures, integrity, and the mass-count distinction. *Synthese* 116(1). 75–111. <http://dx.doi.org/10.1023/a:1005046308299>.
- Moltmann, Friederike. 2004. The semantics of *together*. *Natural Language Semantics* 12(4). 289–318. <http://www.jstor.org/stable/23749465>.
- Moltmann, Friederike. 2009. Degree structure as trope structure: a trope-based analysis of positive and comparative adjectives. *Linguistics and Philosophy* 32(1). 51–94. <http://dx.doi.org/10.1007/s10988-009-9054-5>.
- Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Jaakko Hintikka, Julius Moravcsik & Patrick Suppes (eds), *Approaches to natural language: 1970 Stanford workshop on grammar and semantics*, 221–42. Dordrecht: Springer. http://dx.doi.org/10.1007/978-94-010-2506-5_10.
- Mourelatos, Alexander P. D. 1978. Events, processes, and states. *Linguistics and Philosophy* 2(3). 415–34. <http://dx.doi.org/10.1007/bfoo149015>.
- Müller, Ana & Esmeralda Negrão. 2012. On distributivity in Karitiana. In Patricia Cabredo Hofherr & Brenda Laca (eds), *Verbal plurality and distributivity*, 159–84. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110293500.159>.
- Muskens, Reinhard. 1996. *Meaning and partiality*. Stanford, CA: CSLI Publications.
- Nakanishi, Kimiko. 2004. *Domains of measurement: formal properties of non-split/split quantifier constructions*. Philadelphia, PA: University of Pennsylvania PhD thesis. <http://repository.upenn.edu/dissertations/AAI3138060/>.
- Nakanishi, Kimiko & Maribel Romero. 2004. Two constructions with *most* and their semantic properties. *34th Annual Meeting of the North East Linguistic Society (NELS 34)*. 453–67.
- Naumann, Ralf. 1999. A dynamic logic of events and states for the interaction between plural quantification and verb aspect in natural language. *Logic Journal of the IGPL* 7(5). 591–627. <http://dx.doi.org/10.1093/jigpal/7.5.591>.
- Nouwen, Rick. 2006. *Natural language quantifiers*. Lecture notes for ESSLI 2006 course.

- Nouwén, Rick. 2016. Plurality. In Maria Aloni & Paul Dekker (eds), *The Cambridge Handbook of Formal Semantics*, 267–84. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139236157.010>.
- Oh, Sei-Rang. 2001. Distributivity in an event semantics. *11th Semantics and Linguistic Theory Conference (SALT 11)*, 326–45. <http://dx.doi.org/10.3765/salt.v11i0.2844>.
- Oh, Sei-Rang. 2006. *Plurality markers across languages*. Storrs: University of Connecticut PhD thesis. <http://digitalcommons.uconn.edu/dissertations/AAI3236140>.
- Ojeda, Almerindo E. 2003. A Russellian interpretation of measure nouns. In *From NP to DP: The syntax and semantics of noun phrases*, 255–76. Amsterdam: Benjamins. <http://dx.doi.org/10.1075/la.55.12oje>.
- Parsons, Terence. 1970. An analysis of mass and amount terms. *Foundations of Language* 6, 362–88.
- Parsons, Terence. 1987. Underlying states in the semantical analysis of English. In *Aristotelian Society* 88, 1–30. <http://www.jstor.org/stable/4545070>.
- Parsons, Terence. 1990. *Events in the semantics of English*. Cambridge: MIT Press.
- Parsons, Terence. 1995. Thematic relations and arguments. *Linguistic Inquiry* 26(4), 635–62. <http://www.jstor.org/stable/4178917>.
- Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk, Dick de Jongh & Martin Stokhof (eds), *Studies in Discourse Representation Theory and the theory of generalized quantifiers*, 115–43. Dordrecht: Foris. <http://dx.doi.org/10.1002/9780470758335.ch15>.
- Partee, Barbara H., Alice G. B. ter Meulen & Robert Eugene Wall. 1990. *Mathematical methods in linguistics*. Dordrecht: Kluwer. <http://dx.doi.org/10.1007/978-94-009-2213-6>.
- Partee, Barbara H. & Mats Rooth. 1983. Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds), *Meaning, use and interpretation of language*, 361–83. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110852820.361>.
- Pearson, Hazel. 2009. Towards a semantics for group nouns. MS. Harvard University.
- Pearson, Hazel. 2010. A new semantics for group nouns. *28th West Coast Conference on Formal Linguistics (WCCFL 28)*, 160–8. <http://www.lingref.com/cpp/wccfl/28/paper2448.pdf>.
- Pelletier, Francis Jeffry & Lenhart K. Schubert. 2002. Mass expressions. In Dov M. Gabbay & Franz Guenther (eds), *Handbook of philosophical logic*, vol. 10, 1–87. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-94-017-4524-6_6.
- Petrova, Milena. 2000. The distributive marker *po* in Bulgarian. MS. University of Southern California.
- Pianesi, Fabio. 2002. Book review: Friederike Moltmann, *Parts and Wholes in Semantics*. *Linguistics and Philosophy* 25(1), 97–120. <http://dx.doi.org/10.1023/a:1014347214919>.
- Pianesi, Fabio & Achille C. Varzi. 1996. Events, topology, and temporal relations. *The Monist* 78(1), 89–116. <http://dx.doi.org/10.5840/monist19967919>.
- Pietroski, Paul M. 2005. *Events and semantic architecture*. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199244300.001.0001>.
- Pietroski, Paul M. 2006. Interpreting concatenation and concatenates. *Philosophical Issues* 16(1), 221–45. <http://dx.doi.org/10.1111/j.1533-6077.2006.00111.x>.
- Piñón, Christopher. 1995. *An ontology for event semantics*. Stanford, CA: Stanford University PhD thesis.

- Piñón, Christopher. 1999a. A semantics for durative adverbials. In *Sinn und Bedeutung* 4. Düsseldorf: University of Düsseldorf. http://pinon.sdf-eu.org/work/pinon_sda_ho.pdf.
- Piñón, Christopher. 1999b. Durative adverbials for result states. *18th West Coast Conference on Formal Linguistics (WCCFL 18)*. 420–33. http://pinon.sdfeui.org/work/pinon_dars.pdf.
- Piñón, Christopher. 2000. *Distributive po- in Polish*. Presentation at *Sinn und Bedeutung* 5, Amsterdam. http://pinon.sdfeui.org/work/pinon_dpp_ho_2000.pdf.
- Piñón, Christopher. 2015. An illusory common core. *Theoretical Linguistics* 41(3–4). 185–95. <http://dx.doi.org/10.1515/tl-2015-0012>.
- Pontow, Carsten & Rainer Schubert. 2006. A mathematical analysis of theories of parthood. *Data and Knowledge Engineering* 59(1). 107–38. <http://dx.doi.org/10.1016/j.datak.2005.07.010>.
- Postal, Paul Martin. 1974. *On raising*. Cambridge: MIT Press.
- Potts, Christopher. 2008. Interpretive economy, Schelling points, and evolutionary stability. MS. University of Massachusetts, Amherst. <http://semanticsarchive.net/Archive/jExYWZIN/>.
- Pratt, Ian & Nissim Francez. 2001. Temporal prepositions and temporal generalized quantifiers. *Linguistics and Philosophy* 24(2). 187–255. <http://dx.doi.org/10.1023/a:1005632801858>.
- Przepiórkowski, Adam. 2013. The syntax of distance distributivity in Polish: preserving generalisations with weak heads. *HPSG [Head-driven Phrase Structure Grammar] 2013 Conference*, 161–81. <http://www.stanford.edu/group/cslipublications/cslipublications/HPSG/2013/toc.shtml>.
- Przepiórkowski, Adam. 2014a. Distance distributivity in Polish: towards a Glue Semantics approach. *Empirical Issues in Syntax and Semantics 10 (CSSP [Colloque de Syntaxe et Sémantique à Paris] 2013 Proceedings)*, 107–24. <http://www.cssp.cnrs.fr/eiss10/>.
- Przepiórkowski, Adam. 2014b. Syntactic and semantic constraints in a Glue Semantics approach to distance distributivity. *LFG '14 [Lexical Functional Grammar 2014] Conference*, 482–502. <http://cslipublications.stanford.edu/LFG/19/lfg14.html>.
- Przepiórkowski, Adam. 2015. A weakly compositional analysis of distance distributivity in Polish. *Annual Workshop on Formal Approaches to Slavic Linguistics: The First Berkeley Meeting 2014 (FASL 2014)*. 262–81.
- Quine, Willard van Orman. 1960. *Word and object*. Cambridge: MIT Press.
- Rathert, Monika. 2004. *Textures of time*. Berlin: Akademie. <http://dx.doi.org/10.1515/9783050084015>.
- Richards, Barry. 1982. Tense, aspect and time adverbials, part I. *Linguistics and Philosophy* 5(1). 59–107. <http://dx.doi.org/10.1007/bf00390693>.
- Rimell, Laura. 2004. Habitual sentences and generic quantification. *23rd West Coast Conference on Formal Linguistics (WCCFL 23)*. 663–76. http://www.nyu.edu/projects/szabolcsi/Rimell_WCCFL23_Habitual.pdf.
- Robaldo, Livio. 2011. Distributivity, collectivity, and cumulativity in terms of (in)dependence and maximality. *Journal of Logic, Language and Information* 20(2). 233–71. <http://dx.doi.org/10.1007/s10849-010-9131-8>.
- Roberts, Craige. 1987. *Modal subordination, anaphora, and distributivity*. Amherst: University of Massachusetts PhD thesis. <http://scholarworks.umass.edu/dissertations/AAI8710499/>.
- van Rooij, Robert. 2003. Questioning to resolve decision problems. *Linguistics and Philosophy* 26(6). 727–63. <http://dx.doi.org/10.1023/b:ling.0000004548.98658.8f>.
- van Rooij, Robert. 2008. Comparatives and quantifiers. In *Empirical issues in syntax and semantics*, vol. 7, 393–414. Paris: Colloque de Syntaxe et Sémantique à Paris.

- van Rooij, Robert & Katrin Schulz. 2006. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* 29(2). 205–50. <http://dx.doi.org/10.1007/s10988-005-3760-4>.
- Rothstein, Susan. 1995. Adverbial quantification over events. *Natural Language Semantics* 3(1). 1–31. <http://dx.doi.org/10.1007/bfo1252883>.
- Rothstein, Susan. 2009. Individuating and measure readings of classifier constructions: Evidence from Modern Hebrew. *Brill's Annual of Afroasiatic Languages and Linguistics* 1(1). 106–45. <http://dx.doi.org/10.1163/187666309x12491131130783>.
- Rothstein, Susan. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27(3). 343–97. <http://dx.doi.org/10.1093/jos/ffq007>.
- Sadock, Jerry. 1984. Whither radical pragmatics? In Deborah Schiffrin (ed.), *Meaning, form and use in context: linguistic applications*, 139–49. Washington, DC: Georgetown University Press.
- Safir, Ken & Tim Stowell. 1988. Binominal *each*. *18th Annual Meeting of the North East Linguistic Society (NELS 18)*. 426–50. http://www.linguistics.ucla.edu/people/stowell/Safir-Stowell-_Binominal_Each_-_NELS_1988_ocr.pdf.
- Sakaguchi, Mari. 1998. *Distributivity in English and Japanese*. Los Angeles, CA: University of California PhD thesis.
- Sauerland, Uli. 2003. A new semantics for number. *13th Semantics and Linguistic Theory Conference (SALT 13)*. 258–75. <http://dx.doi.org/10.3765/salt.v13io.2898>.
- Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kepser & Marga Reis (eds), *Linguistic evidence: empirical, theoretical and computational perspectives*, 413–34. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110197549.413>.
- Scha, Remko. 1981. Distributive, collective and cumulative quantification. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds), *Formal methods in the study of language*, 483–512. Amsterdam: Mathematical Center Tracts. <http://dx.doi.org/10.1515/9783110867602.131>. Repr. in Groenendijk, Janssen, and Stokhof (eds), *Truth, interpretation, and information* (Dordrecht: Foris, 1984).
- Schein, Barry. 1986. *Event logic and the interpretation of plurals*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis. <http://hdl.handle.net/1721.1/15107>.
- Schein, Barry. 1993. *Plurals and events*. Cambridge, MA: MIT Press.
- Schlenker, Philippe. 2016. The semantics–pragmatics interface. In Maria Aloni & Paul Dekker (eds), *The Cambridge Handbook of Formal Semantics*, 664–727. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139236157.023>.
- Schwarz, Florian. 2009. *Two types of definites in natural language*. Amherst: University of Massachusetts PhD thesis. http://scholarworks.umass.edu/open_access_dissertations/122/.
- Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer. <http://dx.doi.org/10.1007/978-94-017-2704-4>.
- Schwarzschild, Roger. 2002. The grammar of measurement. *12th Semantics and Linguistic Theory Conference (SALT 12)*. 225–45. <http://dx.doi.org/10.3765/salt.v12io.2870>.
- Schwarzschild, Roger. 2006. The role of dimensions in the syntax of noun phrases. *Syntax* 9(1). 67–110. <http://dx.doi.org/10.1111/j.1467-9612.2006.00083.x>.
- Schwarzschild, Roger. 2014. Distributivity, negation and quantification in event semantics: recent work by L. Champollion. MS. Massachusetts Institute of Technology.
- Schwarzschild, Roger. 2015. Partitives and duratives. *Theoretical Linguistics* 41(3–4). 197–209. <http://dx.doi.org/10.1515/tl-2015-0013>.

- Scontras, Gregory. 2014. *The semantics of measurement*. Cambridge, MA: Harvard University PhD thesis. http://web.stanford.edu/~scontras/papers/Scontras_Thesis.pdf.
- Selkirk, Lisa. 1977. Some remarks on noun phrase structure. In Peter Culicover, Tom Wasow & Adrian Akmajian (eds), *Formal syntax*, 285–316. New York: Academic Press.
- Simons, Peter. 1987. *Parts: a study in ontology*. Oxford: Clarendon Press.
- Singh, Raj. 2007. Presupposition and selectional restrictions. MS. Massachusetts Institute of Technology. <http://http-server.carleton.ca/~singhr/files/singh-selectional-restrictions.pdf>.
- Smith, Carlota. 1997. *The parameter of aspect*, 2nd edn. Dordrecht: Reidel.
- Smollett, Rebecca. 2001. Quantized direct objects don't delimit after all. In Henk J. Verkuyl, Henriëtte de Swart & Angeliek van Hout (eds), *Perspectives on aspect*, 41–59. Berlin: Springer. http://dx.doi.org/10.1007/1-4020-3232-3_3.
- Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: on higher-order implicatures. In Uli Sauerland & Penka Stateva (eds), *Presuppositions and implicature in compositional semantics*, 243–81. London: Palgrave. <http://dx.doi.org/10.1057/9780230210752>.
- Spelke, Elizabeth S. 1990. Principles of object perception. *Cognitive Science* 14(1). 29–56. http://dx.doi.org/10.1207/s15516709cog1401_3.
- Spevak, Olga. 2014. *The noun phrase in classical Latin prose*. Leiden: Brill. <http://dx.doi.org/10.1163/9789004265684>.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19(3). 425–49. <http://www.jstor.org/stable/25164903>.
- Stanley, Jason & Zoltán Gendler Szabó. 2000. On quantifier domain restriction. *Mind and Language* 15(2–3). 219–61. <http://dx.doi.org/10.1111/1468-0017.00130>.
- von Stechow, Arnim. 1984a. Comparing semantic theories of comparison. *Journal of Semantics* 3(1–2). 1–77. <http://dx.doi.org/10.1093/jos/3.1-2.1>.
- von Stechow, Arnim. 1984b. My reaction to Cresswell's, Hellan's, Hoeksema's, and Seuren's comments. *Journal of Semantics* 3(1–2). 183–99. <http://dx.doi.org/10.1093/jos/3.1-2.183>.
- von Stechow, Arnim. 2002a. German *seit* 'since' and the ambiguity of the German Perfect. In Ingrid Kaufmann & Barbara Stiebels (eds), *More than words: a Festschrift for Dieter Wunderlich*, 393–432. Berlin: Akademie. <http://dx.doi.org/10.1515/9783050081274-016>.
- von Stechow, Arnim. 2002b. Temporal prepositional phrases with quantifiers: some additions to Pratt and Francez (2001). *Linguistics and Philosophy* 25(5–6). 755–800. <http://dx.doi.org/10.1023/a:1020872017811>.
- von Stechow, Arnim. 2009. Tenses in compositional semantics. In Wolfgang Klein & Ping Li (eds), *The expression of time*, 129–66. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110199031.129>.
- Sternefeld, Wolfgang. 1998. Reciprocity and cumulative predication. *Natural Language Semantics* 6(3). 303–37. <http://dx.doi.org/10.1023/a:1008352502939>.
- Stickney, Helen. 2008. *The emergence of DP in the partitive structure*. Amherst: University of Massachusetts PhD thesis. http://scholarworks.umass.edu/open_access_dissertations/131/.
- Sutton, Melody. 1993. Binominal "each." Los Angeles: University of California MA thesis.
- Syrett, Kristen. 2013. The role of cardinality in the interpretation of measurement expressions. *Language Acquisition* 20(3). 228–40. <http://dx.doi.org/10.1080/10489223.2013.796949>.
- Syrett, Kristen. 2015. Events and agents in the acquisition of universal quantification. *Theoretical Linguistics* 41(3–4). 211–22. <http://dx.doi.org/10.1515/tl-2015-0014>.

- Szabolcsi, Anna. 2010. *Quantification*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/cbo9780511781681>.
- Szabolcsi, Anna & Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1(3). 235–84. http://dx.doi.org/10.1007/978-94-011-5814-5_7.
- Talmy, Leonard. 1996. Fictive motion in language and “ception.” *Language and Space* 21: 1–276.
- Talmy, Leonard. 2011. Cognitive semantics: an overview. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds), *Semantics: an international handbook of natural language meaning*, 622–42. Berlin: de Gruyter. <http://dx.doi.org/10.1515/9783110226614.622>.
- Tarski, Alfred. 1929. Les fondements de la géométrie des corps (Foundations of the geometry of solids). *Księga Pamiątkowa Pierwszego Polskiego Zjazdu Matematycznego*, suppl. to *Annales de la Société Polonaise de Mathématique* 7. 29–33.
- Tarski, Alfred. 1935. Zur Grundlegung der Booleschen Algebra. *Fundamenta Mathematicae* 24, 177–98. Translated into English as ‘On the foundations of the Boolean algebra’ (1956). <https://eudml.org/doc/212745>.
- Tarski, Alfred. 1956. On the foundations of the Boolean algebra. In Joseph Henry Woodger (ed.), *Logics, semantics, metamathematics: papers from 1923 to 1938*, 320–41. Oxford: Clarendon Press.
- Tasmowski-De Ryck, Liliane & S. Paul Verluyten. 1982. Linguistic control of pronouns. *Journal of Semantics* 1(4). 323–46. <http://dx.doi.org/10.1093/jos/1.3-4.323>.
- Tatevosov, Sergej. 2002. The parameter of actionality. *Linguistic Typology* 6(3). 317–401. <http://dx.doi.org/10.1515/lity.2003.003>.
- Taub, Alison. 1989. Collective predicates, aktionsarten and *all*. In Emmon Bach, Angelika Kratzer & Barbara H. Partee (eds), *Papers on quantification*, 337–66. Amherst: University of Massachusetts.
- Taylor, Barry. 1977. Tense and continuity. *Linguistics and Philosophy* 1(2). 199–220. <http://dx.doi.org/10.1007/bf00351103>.
- Taylor, Barry. 1985. *Modes of occurrence*. Oxford: Blackwell.
- Tellier, Christine & Daniel Valois. 1993. Binominal *chacun* and pseudo-opacity. *Linguistic Inquiry* 24(3). 399–589. <http://www.jstor.org/stable/4178830>.
- Thomas, Guillaume. 2015. The present tense is not vacuous. *Journal of Semantics* 32(4). 685–747. <http://dx.doi.org/10.1093/jos/ffu010>.
- Thomas, Guillaume & Yasutada Sudo. 2016. Cumulative readings of *each*. Presentation at the Workshop on (Co-)Distributivity, Paris, February 11–12.
- Todorova, Marina, Karen Straub, William Badecker & Robert Frank. 2000. Aspectual coercion and the online computation of sentential aspect. In Lila R. Gleitman & Aravind K. Joshi (eds), *22nd Annual Conference of the Cognitive Science Society*, 3–8. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.23.8886&rep=rep1&type=pdf>.
- Tsoulas, George & Eytan Zweig. 2009. *Distributivity across languages and logics*. Course at ESSLLI. <http://distshare.blogspot.com>.
- Tunstall, Susanne Lynn. 1998. *The interpretation of quantifiers: semantics and processing*. Amherst: University of Massachusetts PhD thesis. <http://scholarworks.umass.edu/dissertations/AAI9909228/>.
- Ursini, Francesco-Alessio. 2006. *Measure phrases and directional PPs: measuring out movement*. Utrecht: Utrecht University MA thesis. <http://dspace.library.uu.nl/handle/1874/13816>.

- Vaillette, Nathan. 2001. Flexible summativity: a type-logical approach to plural semantics. In *OSU Working Papers in Linguistics* 56, 135–57. Columbus: Ohio State University. <http://linguistics.osu.edu/sites/linguistics.osu.edu/files/osuwpl56-nva.pdf>.
- Varzi, Achille C. 2006. A note on the transitivity of parthood. *Applied Ontology* 1(2). 141–6. http://academiccommons.columbia.edu/download/fedora_content/download/ac:179023/CO-NTENT/AO_2006.pdf.
- Varzi, Achille C. 2016. Mereology. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2016. Stanford, CA. <http://plato.stanford.edu/archives/win2016/entries/mereology/>.
- Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 66(2). 143–60. <http://dx.doi.org/10.2307/2182371>.
- Vendler, Zeno. 1962. *Each and every, any and all*. *Mind* 71(282). 145–60. <http://dx.doi.org/10.1093/mind/lxxi.282.145>.
- Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
- Verkuyl, Henk J. 1972. *On the compositional nature of the aspects*. Dordrecht: Reidel. <http://dx.doi.org/10.1007/978-94-017-2478-4>.
- Verkuyl, Henk J. 1989. Aspectual classes and aspectual composition. *Linguistics and Philosophy* 12(1). 39–94. <http://dx.doi.org/10.1007/bf00627398>.
- Verkuyl, Henk J. 1993. *A theory of aspectuality: the interaction between temporal and atemporal structure*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/cbo9780511597848>.
- Verkuyl, Henk J. 2005. Aspectual composition: surveying the ingredients. In Henk J. Verkuyl, Henriëtte de Swart & Angeliek van Hout (eds), *Perspectives on aspect*, 19–39. Berlin: Springer. http://dx.doi.org/10.1007/1-4020-3232-3_2.
- Verkuyl, Henk J. & Jaap van der Does. 1991. *The semantics of plural noun phrases*. Amsterdam: University of Amsterdam. <https://www.illc.uva.nl/Research/Publications/Reports/LP-1991-07.text.pdf>.
- Vlach, Frank. 1993. Temporal adverbials, tenses and the Perfect. *Linguistics and Philosophy* 16(3). 231–83. <http://dx.doi.org/10.1007/bf00985970>.
- de Vries, Hanna. 2015. *Shifting sets, hidden atoms: the semantics of distributivity, plurality and animacy*. Utrecht University PhD thesis. <http://dspace.library.uu.nl/handle/1874/312186>.
- Wechsler, Stephen & Yae-Sheik Lee. 1996. The domain of direct case assignment. *Natural Language and Linguistic Theory* 14(3). 629–64. <http://dx.doi.org/10.1007/bf00133600>.
- Williams, Alexander. 2009. Themes, cumulativity, and resultatives: comments on Kratzer 2003. *Linguistic Inquiry* 40(4). 686–700. <http://dx.doi.org/10.1162/ling.2009.40.4.686>.
- Williams, Edwin. 1991. Reciprocal scope. *Linguistic Inquiry* 22(1). 159–73. <http://www.jstor.org/stable/4178711>.
- Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20(4). 399–467. <http://dx.doi.org/10.1023/a:1005354323136>.
- Winter, Yoad. 2001. *Flexibility principles in Boolean semantics: the interpretation of coordination, plurality, and scope in natural language*. Cambridge, MA: MIT Press.
- Winter, Yoad. 2002. Atoms and sets: a characterization of semantic number. *Linguistic Inquiry* 33(3). 493–505. <http://dx.doi.org/10.1162/002438902760168581>.

- Winter, Yoad & Remko Scha. 2015. Plurals. In Shalom Lappin & Chris Fox (eds), *Handbook of contemporary semantic theory*, 2nd edn, 77–113. Oxford: Wiley-Blackwell. <http://dx.doi.org/10.1002/9781118882139.ch3>.
- Zimmermann, Malte. 2002a. A compositional analysis of anti-quantifiers as quantifiers. *12th Semantics and Linguistic Theory Conference (SALT 12)*. 322–38. <http://dx.doi.org/10.3765/salt.v12i0.2876>.
- Zimmermann, Malte. 2002b. *Boys buying two sausages each: on the syntax and semantics of distance-distributivity*. Amsterdam: University of Amsterdam PhD thesis. <http://hdl.handle.net/11245/1.194584>.
- Zimmermann, Malte. 2008. Quantification in Hausa. In Lisa Matthewson (ed.), *Quantification: universals and variation*, 415–75. Bingley, UK: Emerald.
- Zucchi, Sandro & Michael White. 2001. Twigs, sequences and the temporal constitution of predicates. *Linguistics and Philosophy* 24(2). 223–70. <http://dx.doi.org/10.1023/a:1005690022190>.
- Zwarts, Joost. 1997. Vectors as relative positions: a compositional semantics of modified PPs. *Journal of Semantics* 14(1). 57–86. <http://dx.doi.org/10.1093/jos/14.1.57>.
- Zwarts, Joost. 2005. Prepositional aspect and the algebra of paths. *Linguistics and Philosophy* 28(6). 699–740. <http://dx.doi.org/10.1007/s10988-005-2466-y>.
- Zwarts, Joost. 2006. Event shape: paths in the semantics of verbs. MS. Radboud University Nijmegen and Utrecht University. <http://www.let.uu.nl/users/Joost.Zwarts/personal/EventShape.pdf>.
- Zwarts, Joost. 2013. From N to N: the anatomy of a construction. *Linguistics and Philosophy* 36(1). 65–90. <http://dx.doi.org/10.1007/s10988-013-9131-7>.
- Zwarts, Joost & Henk J. Verkuyl. 1994. An algebra of conceptual structure: an investigation into Jackendoff's conceptual semantics. *Linguistics and Philosophy* 17(1). 1–28. <http://dx.doi.org/10.1007/bf00985039>.
- Zweig, Eytan. 2008. *Dependent plurals and plural meaning*. New York: New York University PhD thesis. <http://semanticsarchive.net/Archive/GJhMzhZ/>.
- Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32(4). 353–407. <http://dx.doi.org/10.1007/s10988-009-9064-3>.

Index of languages

Albanian 201–2, 210

Bulgarian 205–6, 210, 218–19, 222

Chinese 237

Czech 205–6, 210, 218–19, 222

Dutch 55, 201–2, 210, 218, 222

Finnish 278

French 201–2, 210, 218, 222

German 8, 65, 158, 169, 180, 183, 192–3,
197–201, 205, 209–10, 213, 216, 218–22,
230–1, 237, 274, 277

Gujarati 277, 278

Hausa 208

Hindi 278

Hungarian 208, 278

Icelandic 201–3, 210, 222

Italian 201, 203, 210, 218, 222

Japanese 201, 203, 205, 210, 277

Karitiana 208

Korean 205–6, 210, 222, 278

Latin 201, 204–5, 209–10

Norwegian 201, 203–4, 210, 218, 222

Polish 205–7, 210, 222

Romanian 205, 207, 210

Russian 201, 204–5, 207, 210, 218, 222

Telugu 208

Tlingit 208–9

Turkish 278

Index of people

- Abusch, Dorit 66, 105–6
 Akmajian, Adrian 63
 Aldridge, Maurice Vincent 225
 Anagnostopoulou, Elena 105
- Bach, Emmon 1, 24, 27–8, 70, 89
 Badecker, William 189
 Bale, Alan 63, 152, 278
 Balusu, Rahul 198, 208
 Barker, Chris 47–8, 73, 75, 83, 196
 Barner, David 26, 103
 Barwise, Jon 41, 224
 Bayer, Samuel L. 51, 171, 227
 Bealer, George 49
 Beaver, David 128
 Beavers, John 123
 Beck, Sigrid 165, 181, 191
 Beghelli, Filippo 224–5
 Bennett, Michael R. 5, 18, 47, 87, 102, 122–3
 van Benthem, Johan 24, 29
 Binnick, Robert 65
 Bittner, Marian 234
 Blaheta, Don 77, 160, 233, 235
 Bobaljik, Jonathan D. 197, 212
 Borik, Olga 109–10, 116, 204
 Bortolussi, Bernard 204–5
 Brasoveanu, Adrian 21, 41–2, 57, 207, 224, 243, 260
 Breheny, Richard 41
 Brisson, Christine 163, 241–2, 244–5, 247–8, 255, 258, 265
 Bunt, Harry 18, 26, 42, 103, 251
 Büring, Daniel 197, 209, 221
 Burzio, Luigi 196–7, 203, 209, 220
- Cable, Seth 197, 208–9, 213
 Carlson, Gregory N. 26–7, 34, 49–51, 53, 58, 60, 105, 179–80, 277
- Carlson, Lauri 251
 Carnie, Andrew 34
 Carston, Robyn 41
 Cartwright, Helen 30
 Casati, Roberto 13, 18, 43
 Chierchia, Gennaro 26, 42–4, 46, 49–50, 63, 103, 142, 277
 Choe, Jae-Woong 72, 77, 196, 205–6
 Chomsky, Noam 33–4, 110, 197, 209, 278
 Clarke, T. H. 147
 Cohen, Ariel 45
 Collins, Chris 231
 Comrie, Bernard 117
 Cooper, Robin 41, 224
 Crain, Stephen 279
 Cresswell, Max 18, 24, 30–1
 Cvitaš, Tomislav 141
- Daubert, B. F. 147
 Davidson, Donald 27, 29, 37, 50, 146
 de Mey, Sjaap 45
 de Swart, Henriëtte 43–4, 278
 de Vries, Hanna 73
 Dekker, Paul 134
 Delfitto, Denis 49
 Deo, Ashwini Sharad 66, 157, 165, 179, 181, 183, 189–90, 215, 277–8
 Dobrovie-Sorin, Carmen 245, 248, 250–1, 253, 275
 van der Does, Jaap 163, 169
 Doetjes, Jenny 26, 102, 212, 275–6
 Dotlačil, Jakub 211, 213, 224, 233
 Dougherty, Ray C. 243
 Dowty, David R. 7–8, 33, 39, 65–6, 73, 86–8, 100–10 *passim*, 117–29 *passim*, 135, 143, 157, 162, 180, 182, 211–12, 241–8 *passim*, 254, 266, 268, 271, 274, 277

- Eberle, Kurt 123, 261, 276
 Eckardt, Regine 27, 28, 177–8
 Egg, Markus 54, 181, 276
 Erlewine, Michael Yoshitaka 112, 267
- Farkas, Donka F. 43–4, 207–8, 225, 228, 278
 Ferreira, Marcelo 181, 225–6, 278
 Filip, Hana 123, 206, 278
 Fine, Kit 178
 Fox, Danny 30, 32
 Francez, Nissim 191
 Frank, Robert 189
- Gagnon, Michael 278
 Gärdenfors, Peter 279
 Garey, Howard B. 117
 Gawron, Jean Mark 7, 63, 117–20, 123, 272
 van Geenhoven, Veerle 54, 104, 158, 179
 Geurts, Bart 41
 Gil, David 77, 198, 207–8
 Gillon, Brendan S. 103, 111, 156, 163, 166, 191, 194, 200
 Graham, Daniel W. 117
 Grevisse, Maurice 202
 Gruber, Jeffrey S. 33
 Guéron, Jacqueline 63
- Hackl, Martin 30, 32, 41, 243, 245
 Harnish, Robert M. 41
 Heim, Irene 8, 10, 12, 60, 112, 168, 211
 Hellan, Lars 24, 30
 Hendriks, Herman 196
 Heny, Frank 105
 Higginbotham, James 251
 Hinrichs, Erhard 12, 29, 36, 58, 101, 105–7, 110, 123, 125
 Hitzeman, Janet 105, 277
 Hoeksema, Jack 81
 Hogg, Richard M. 225
 Homann, Klaus 141
 Horn, Laurence 41
 Hovda, Peter 14, 16, 18
- Iatridou, Sabine 105
 Ionin, Tania 50, 278
 Iordăchioaia, Gianina 207
 Ioup, Georgette 224
 Izvorski, Roumyana 105
- Jackendoff, Ray S. 33, 221, 279
 Janke, Vikki 217
 Jayaseelan, Karattuparambil A. 198, 208
 Junker, Marie-Odile 202
- Kallay, Nikola 141
 Kamp, Hans 24, 105
 Karttunen, Lauri 96, 277
 Keim, Marybelle C. 221
 Keizer, Evelien 62
 Kennedy, Christopher 30, 41, 169
 Khanjian, Hrayr 278
 Kiparsky, Paul 278
 Kipper-Schuler, Karin 33
 Kleiber, Georges 73
 Klein, Ewan 30
 Kotek, Hadas 112, 267
 Krantz, David H. 141
 Krasikova, Sveta 30
 Kratzer, Angelika 8, 10, 12, 24, 26, 33, 35, 37, 39, 43, 51–4, 58, 60, 114, 125, 133–4, 156, 165, 178–84, 192, 225–8, 236, 274, 277
 Krifka, Manfred 2–58 *passim*, 73, 89, 104, 106, 109, 115–53 *passim*, 165, 182–3, 245, 261, 270–2, 275, 277–9
 Križ, Manuel 73, 242
 Kroch, Anthony S. 56, 73, 224, 243–5, 253–4
 Kuchitsu, Kozo 141
 Kuhn, Jeremy 197, 245–58 *passim*, 275–6
 Kuno, Susumu 77
- Ladusaw, William 254, 278
 Landman, Fred 6, 10, 12, 17–18, 22, 24, 28, 34, 41, 52, 54–8, 63, 66, 73, 81, 104, 155, 167, 181, 200, 229–30, 232, 254, 261, 264, 270, 276
 Langacker, Ronald 4

- Lasersohn, Peter 32, 36, 50, 52, 54, 81, 155–78
passim, 191–5, 220, 242, 266, 273
- Lassiter, Daniel 30
- LaTerza, Chris 176, 196, 213, 215, 237
- Lee, Yae-Sheik 278
- Lehrer, Adrienne 63
- Levin, Beth 33
- Levinson, Stephen C. 41
- Lewis, David 30, 196
- Lin, Jo-Wang 237
- Link, Godehard 1–2, 6–7, 10, 12, 17–19, 24–5,
36, 44, 49–50, 53, 68, 70, 77, 101, 108–11,
149, 155–75 *passim*, 191, 194–5, 209–10,
231–2, 237, 241, 245, 270, 273–4, 277
- Lønning, Jan Tore 31–2, 37–9, 142, 251, 270
- Luce, R. Duncan 141
- Malamud, Sophia A. 163, 167, 242
- Maling, Joan M. 217
- Marantz, Alec 33
- Markosian, Ned 27
- Matthewson, Lisa 225, 228, 241, 251, 278
- Matushansky, Ora 50, 278
- McCawley, James D. 96
- ter Meulen, Alice 49–50, 277
- de Mey, Sjaap 45
- Mills, Ian 141
- Milsark, Gary 49
- Mittwoch, Anita 105, 276
- Mollá-Aliod, Diego 104
- Moltmann, Friederike 13–14, 37, 101, 107–8,
110, 112, 118, 122–3, 125–6, 144, 157, 182–3,
197–8, 211, 218, 243, 245, 266, 272
- Montague, Richard 12, 220, 224
- Mourelatos, Alexander P. D. 27–8
- Müller, Ana 208
- Murray, John P. 221
- Muskens, Reinhard 96
- Nakanishi, Kimiko 242, 277
- Naumann, Ralf 276
- Neeleman, Ad 217
- Negrão, Esmeralda 208
- Nouwen, Rick 41, 155
- Oh, Sei-Rang 195–6, 198, 205
- Ojeda, Almerindo E. 30
- Parsons, Terence 27–8, 30, 33–4, 51, 58,
105, 191
- Partee, Barbara H. 5, 10, 41, 66, 69, 87, 102,
122–3, 244
- Pearson, Hazel 27, 46, 48
- Pelletier, Francis Jeffry 42, 49–50, 277
- Peters, Stanley 96
- Petrova, Milena 205–6
- Pianesi, Fabio 13, 24, 108
- Pietroski, Paul M. 60
- Piñango, María Mercedes 66, 157, 165, 179,
181, 183, 189–90, 277–8
- Piñón, Christopher 27, 65–6, 83, 87–8, 95,
104, 112–13, 205, 262, 271–2
- Pontow, Carsten 15–16
- Postal, Paul Martin 196, 231
- Potts, Christopher 169
- Pratt, Ian 191
- Przepiórkowski, Adam 205–7, 222–3, 233
- Quine, Willard van Orman 23, 25, 69, 103,
142–3
- Rathert, Monika 66, 105, 107
- Reyle, Uwe 105
- Richards, Barry 105
- Rimell, Laura 277
- Robaldo, Livio 41
- Roberts, Craige 155, 159–60, 162, 196, 199,
210, 273
- Romero, Maribel 243
- van Rooij, Robert 30, 41, 167
- Rooth, Mats 66, 69, 105–6
- Rothstein, Susan 43, 47–8, 63, 66, 104,
169, 181, 229, 276
- Sadock, Jerry 41
- Sakaguchi, Mari 203, 205
- Sauerland, Uli 44, 165
- Scha, Remko 52, 55, 71, 155

- Schein, Barry 8, 51–2, 55, 171, 176–7, 217, 226–7, 274
- Schlenker, Philippe 188, 264
- Schubert, Lenhart K. 15–16, 42
- Schulz, Katrin 41
- Schwarzschild, Roger 2, 6–7, 18, 37–9, 45–7, 60, 63, 80, 83, 87–8, 90–1, 95, 111–13, 116, 139–75 *passim*, 183, 188, 191–5, 200, 210–11, 215, 236–7, 257, 261–2, 271–9 *passim*
- Scontras, Gregory 63
- Selkirk, Lisa 62, 139
- Simons, Peter 13, 18
- Singh, Raj 96
- Smith, Carlota 27
- Smollett, Rebecca 124
- Snedeker, Jesse 26, 103
- Soare, Elena 207
- Spector, Benjamin 44, 46, 165, 256, 263–4
- Spelke, Elizabeth S. 25
- Spevak, Olga 205
- Sportiche, Dominique 197, 212
- Stanley, Jason 190, 228
- von Stechow, Arnim 29, 106, 110, 175, 181, 191, 277–8
- Sternefeld, Wolfgang 165
- Stickney, Helen 63
- Straub, Karen 189
- Sudo, Yasutada 224, 226
- Suppes, Patrick 141
- Sutton, Melody 213
- de Swart, Henriëtte 43–4, 278
- Syrett, Kristen 279
- Szabó Zoltán Gendler 190, 228
- Szabolcsi, Anna 30, 155, 208, 213, 224
- Talmy, Leonard 119, 279
- Tarski, Alfred 14, 16–18
- Tasmowski-De Ryck, Liliane 230
- Tatevosov, Sergej 118
- Taub, Alison 244–7, 255, 258
- Taylor, Barry 100–3, 171, 226
- Tellier, Christine 202
- ter Meulen, Alice 49–50, 277
- Thomas, Guillaume 224–6
- Todorova, Marina 189
- Tsoulas, George 71
- Tunstall, Susanne Lynn 224–5
- Tversky, Amos 141
- Ursini, Francesco-Alessio 29, 227
- Vaillette, Nathan 18, 20, 112
- Valois, Daniel 202
- van Benthem, Johan 24, 29
- van der Does, Jaap 163, 169
- van Geenhoven, Veerle 54, 104, 158, 179
- van Rooij, Robert 30, 41, 167
- Varzi, Achille C. 13, 15, 18, 24, 43, 108
- Vendler, Zeno 86, 117–18, 122
- Verkuy, Henk 2, 65, 67, 117–18, 163, 180, 276, 279
- Verluyten, S. Paul 230
- Vlach, Frank 66, 104
- von Stechow, Arnim 29, 106, 110, 175, 181, 191, 277–8
- de Vries, Hanna 73
- Wall, Robert Eugene 10
- Wechsler, Stephen 278
- White, Michael 8, 43, 108, 144, 157, 179, 182–3, 228, 248, 274, 276
- Williams, Alexander 13, 35, 51, 236
- Williams, Edwin 211
- Winter, Yoad 8, 47, 69, 72–5, 155–6, 159–63, 166, 191, 220, 242–8, 251, 259, 268, 274
- Zimmermann, Malte 8, 77, 158, 160, 192–209 *passim*, 217–37 *passim*, 274
- Zucchi, Sandro 8, 43, 108, 144, 157–8, 179, 182, 223, 228, 248, 274, 276
- Zwarts, Frans 30
- Zwarts, Joost 29, 36, 46, 54, 57–8, 102, 279
- Zweig, Eytan 8, 45–6, 50, 71, 165, 227, 239–40, 260, 263–4, 266, 274, 276

Index of topics

- adverbial distributive items 193, 211
 - see also *each*; *jewels*
- algebraic closure 19–20, 43–4, 53, 111–12, 176
- algebraic semantics 12, 51
- all* 70, 241–3, 250–65
- antisymmetry, axiom of 13
- anti-vagueness condition 26
- approximate divisive reference 59, 148–9
- aspect
 - spatial 139
 - temporal 117–18
- atelic
 - aspect 70
 - predicate 78, 86–7, 91, 101–2, 106–7
- atom predicates (Winter) 73, 243
- atomic granularity 169–70, 192
- atomicity 19, 21, 26, 40, 57, 170, 229
- atoms (mereological) 21–2, 25–6, 43, 57, 106, 256–7
 - temporal 29, 153, 188
 - group 29, 200
 - impure 22, 48, 57, 266
 - pure 23, 57
- backwards distributivity 220–1
- binding theory 209, 221
- Boolean algebra 16–17
- boundedness 120, 269, 278
 - and unboundedness 1–2, 269, 279
 - question 2–3, 6, 100
- cardinality function 40–1
- c-command 45, 60, 197, 231
- classical extensional mereology (CEM) 6, 8, 10–13, 24, 106, 142
- collective
 - entailment 98, 156, 161
 - predicates 2, 6, 8, 22, 69, 72–4, 84
- collectivity 72, 161, 239
- coordination 63, 67, 72, 74
- count/mass distinction 42, 65, 138
- cover 80, 101, 110–12, 144–5, 163–71, 174–5, 184–6
 - predicate 111, 186
- cumulative reading 8, 20, 55–6, 60, 155, 164, 226–8, 240, 260–7, 274
- cumulativity
 - assumption 34–6, 54, 120, 165, 168, 192
 - lexical c. 52–4, 79, 106, 165, 168, 189–90
 - of thematic roles 79, 227
- D operator 7, 58, 155, 158–78, 193–4, 210–11, 231–2, 257, 268, 280
- degrees 29–33, 37–9, 64
- dependent plurals 8, 43, 45, 240, 264
- determiners
 - quantificational 73
 - plural 73
- dimension parameter 4, 91, 121–2, 125, 145, 149, 158–60, 171, 183, 186, 210, 216, 219–24, 249, 272
- distance-distributive items 8, 51, 194–8, 201, 205, 209–10, 230, 236–7
- distributive operator *see* D operator; Part operator
- distributive
 - polarity items 195
 - predicates 2, 4–5, 55–6, 72–5, 78–85, 100, 161, 240, 249–52, 265, 269
 - readings 3, 58, 164, 209, 236, 267
 - relation 70–2, 76–8
- distributivity 71–5
 - atomic d. 159, 160–3
 - backwards d. 220–1
 - component 45
 - constraint 95, 139

- covert d. 155–60
 event-based nonatomic d. 156–63
 lexical d. 169, 189
 over conjuncts 218
 phrasal d. 155, 161–72, 179, 189
 temporal d. 35, 158
 divisive reference 7, 23, 49, 89–93, 101, 103,
 120–9, 133, 138, 148–9, 251
 divisiveness clause 129, 131, 133
 downward-entailing contexts 45–6, 276
- each* 68–9, 71, 75, 158–60, 217–18, 258–60
 ensemble theory 18
 event identification 58, 60
 events 27–8
 non-punctual 79
 punctual 5, 53, 79
every 71, 197, 224–8
 existence clause 129, 131
 existential closure 58–9, 69, 84, 128, 176, 234
- for*-adverbials 61, 65–8, 75–6, 134–6, 179–91
 functions 32
 cardinality f. 40
 measure f. 37–8, 63, 78, 90–1, 139–53
 trace f. 35–8, 69, 126, 139, 159
 unit f. 38–40, 63
- gather*-type predicates 71, 245–50
 granularity parameter 4–5, 7, 96, 100–1,
 152–3, 159–60, 170–5, 182–90, 216, 229,
 249, 255, 271
 level 94–5, 114, 147, 258–60
- holding at an interval 36, 126
 homomorphism, sum 33–8, 52, 87, 126, 174,
 181, 216, 229, 259
- Interpretive Economy 169–70
 intervals 28–9
 spatial 30–1, 57–8, 126–7, 136
 temporal 23, 31, 39, 104, 108–9, 119, 123,
 135, 156, 171, 210
- jewels* 8, 158, 168–9, 193–201, 218–22
- Key 3, 77–8, 116
 kind/object opposition 50
 kinds 26–7, 49–50
- λ -abstraction 162, 195, 234–5
 index-triggered 234–5
 lexical cumulativity 52–4, 79, 106, 165, 168,
 185–9, 261
 logical form (LF) 69
- Map 3, 92–6
 mass reference 2, 70, 101, 104, 110
 maximizing effect of *all* 241–2
 meaning postulate 7–8, 71, 81–5, 115, 165
 measure function 139
 extensive 7, 141–2
 intensive 139, 143, 151
 monotonic 143
 Quinean 142–3
 mereological sum 142, 145
 mereology 12, 28
 classical extensional m. (CEM) 13–17, 24
 minimal parts problem 5, 87, 91, 94, 101–9
 monotonicity 143–6
 multiplicity component 45
- Neo-Davidsonian semantics 5, 50–1, 191
 nonatomic distributivity 156–163
 nonatomic entities 156, 193–4
 noun phrases 54–8
 nouns
 mass n. 42, 49–50
 fake mass n. 26
 count n. 42–6
 group n. 42, 47–9
 measure n. 42, 46–7, 62
 numbers (type) 31–2, 38
 numerals 41, 57
numerous-type predicates 74, 243–4
- occasion reading of distributive
 items 198–209, 218, 230, 232

- ontology 23–32
- ordinary objects 25
- overlap 14

- part of* construction 14
- Part operator 7, 157, 167–8, 171–5, 185–6, 194–5, 200, 210–11, 215, 232
- parthood 12, 43, 144–5
 - axiom of 13
 - proper p. 79
 - relation 17
 - structured 13
 - unstructured 13
- partitives
 - group p. 48
 - true p. 62, 141, 153, 278
 - see also* pseudopartitives
- Perfect 66
- plurals, dependent 8, 43, 45, 240, 264
- prepositional phrases 29, 57, 217, 233, 277
- pseudopartitives 2, 30, 38–9, 47–8, 50, 62–5, 67–8, 76–8, 89–101, 140–53, 250

- quantifier raising 58, 183
- quantifier
 - existential q. 58, 104, 109–10
 - generalized q. 41, 224

- reflexivity 13–14, 20, 79
- restriction to relevance 134
- runtime 31, 36, 102, 171

- scalar implicature 4, 46, 188
- semilattice, complete join 17
- set predicates 74
- set theory (Winter) 17–19
- Share 3, 63, 70, 77, 84–90, 92–5, 147, 158, 271
- spatial adverbials 119
- star operator *see* algebraic closure
- strata 5
 - theory 1, 4, 50, 70, 101, 145–53, 175, 195, 231, 253, 269
- stratified distributive reference (SDR) 78–80
- stratified measurement reference (SMR) 91–5
- stratified reference (SR) 4, 70, 78, 94
 - restricted 262, 280–1
 - universal 94–5, 249, 262, 281
- stratified subinterval reference (SSR) 87, 88, 94, 111, 113
- subgroup distributivity 8, 248, 250–3
- subinterval property 5, 86–7, 94, 102–9, 123, 178
- sum 12–15, 22
 - arithmetic s. 31
 - binary s. 14
 - generalized s. 16
 - mereological s. 31
 - operation 20
 - pointwise 16
 - proper 18
 - uniqueness of 15
- telic vs. atelic predicates 2, 5, 78–9, 102, 106, 114
- temporal distributivity 35, 158
- thematic roles 27, 32, 33–5, 50–1, 55, 58, 60, 69
 - cumulativity assumption for 34
- trace functions 35–8, 69, 126, 139, 159
- transitivity 13, 107
- type shifters 58–60
- types, semantic 11

- unboundedness 2, 7, 100, 153, 267
- Unique Role Requirement (URR) 34
- unique separation 17, 106, 133
- universal stratified reference 94–5, 249

- v ('little v ') 68

- world knowledge 6, 83, 99, 181

- Zimmermann's generalization 201
- Zweig's observation 240

OXFORD STUDIES IN THEORETICAL LINGUISTICS

Published

- 1 The Syntax of Silence
Sluicing, Islands, and the Theory of Ellipsis
by Jason Merchant
- 2 Questions and Answers in Embedded Contexts
by Utpal Lahiri
- 3 Phonetics, Phonology, and Cognition
edited by Jacques Durand and Bernard Laks
- 4 At the Syntax–Pragmatics Interface
Concept Formation and Verbal Underspecification
in Dynamic Syntax
by Lutz Marten
- 5 The Unaccusativity Puzzle
Explorations of the Syntax–Lexicon Interface
edited by Artemis Alexiadou, Elena
Anagnostopoulou, and Martin Everaert
- 6 Beyond Morphology
Interface Conditions on Word Formation
by Peter Ackema and Ad Neeleman
- 7 The Logic of Conventional Implicatures
by Christopher Potts
- 8 Paradigms of Phonological Theory
edited by Laura Downing, T. Alan Hall, and Renate
Raffelsiefen
- 9 The Verbal Complex in Romance
by Paola Monachesi
- 10 The Syntax of Aspect
Deriving Thematic and Aspectual Interpretation
Edited by Nomi Erteschik-Shir and Tova Rapoport
- 11 Aspects of the Theory of Clitics
by Stephen Anderson
- 12 Canonical Forms in Prosodic Morphology
by Laura J. Downing
- 13 Aspect and Reference Time
by Olga Borik
- 14 Direct Compositionality
edited by Chris Barker and Pauline Jacobson
- 15 A Natural History of Infixation
by Alan C. L. Yu
- 16 Phi-Theory
Phi-Features Across Interfaces and Modules
edited by Daniel Harbour, David Adger, and
Susana Béjar
- 17 French Dislocation
Interpretation, Syntax, Acquisition
by Cécile De Cat
- 18 Inflectional Identity
edited by Asaf Bachrach and Andrew Nevins
- 19 Lexical Plurals
by Paolo Acquaviva
- 20 Adjectives and Adverbs
Syntax, Semantics, and Discourse
edited by Louise McNally and Christopher Kennedy
- 21 InterPhases
Phase-Theoretic Investigations of Linguistic
Interfaces
edited by Kleanthes Grohmann
- 22 Negation in Gapping
by Sophie Repp
- 23 A Derivational Syntax for Information Structure
by Luis López
- 24 Quantification, Definiteness, and
Nominalization
edited by Anastasia Giannakidou and Monika
Rathert
- 25 The Syntax of Sentential Stress
by Arsalan Kahnemuyipour
- 26 Tense, Aspect, and Indexicality
by James Higginbotham
- 27 Lexical Semantics, Syntax, and Event Structure
edited by Malka Rappaport Hovav, Edit Doron, and
Ivy Sichel
- 28 About the Speaker
Towards a Syntax of Indexicality
by Alessandra Giorgi
- 29 The Sound Patterns of Syntax
edited by Nomi Erteschik-Shir and Lisa Rochman
- 30 The Complementizer Phase
edited by Phoevos Panagiotidis
- 31 Interfaces in Linguistics
New Research Perspectives
edited by Raffaella Folli and Christiane Ulbrich
- 32 Negative Indefinites
by Doris Penka
- 33 Events, Phrases, and Questions
by Robert Truswell
- 34 Dissolving Binding Theory
by Johan Rooryck and Guido Vanden Wyngaerd
- 35 The Logic of Pronominal Resumption
by Ash Asudeh
- 36 Modals and Conditionals
by Angelika Kratzer
- 37 The Theta System
Argument Structure at the Interface
edited by Martin Everaert, Marijana Marelj, and
Tal Siloni
- 38 Sluicing
Cross-Linguistic Perspectives
edited by Jason Merchant and Andrew Simpson

- 39 Telicity, Change, and State
A Cross-Categorial View of Event Structure
edited by Violeta Demonte and Louise McNally
- 40 Ways of Structure Building
edited by Myriam Uribe-Etxebarria and Vidal Valmala
- 41 The Morphology and Phonology of Exponence
edited by Jochen Trommer
- 42 Count and Mass Across Languages
edited by Diane Massam
- 43 Genericity
edited by Alda Mari, Claire Beyssade, and Fabio Del Prete
- 44 Strategies of Quantification
edited by Kook-Hee Gil, Steve Harlow, and George Tsoulas
- 45 Nonverbal Predication
Copular Sentences at the Syntax–Semantics Interface
by Isabelle Roy
- 46 Diagnosing Syntax
edited by Lisa Lai-Shen Cheng and Norbert Corver
- 47 Pseudogapping and Ellipsis
by Kirsten Gengel
- 48 Syntax and its Limits
edited by Raffaella Folli, Christina Sevdali, and Robert Truswell
- 49 Phrase Structure and Argument Structure
A Case Study of the Syntax–Semantics Interface
by Terje Lohndal
- 50 Edges in Syntax
Scrambling and Cyclic Linearization
by Heejeong Ko
- 51 The Syntax of Roots and the Roots of Syntax
edited by Artemis Alexiadou, Hagit Borer, and Florian Schäfer
- 52 Causation in Grammatical Structures
edited by Bridget Copley and Fabienne Martin
- 53 Continuations and Natural Language
by Chris Barker and Chung-chieh Shan
- 54 The Semantics of Evaluativity
by Jessica Rett
- 55 External Arguments in Transitivity Alternations
by Artemis Alexiadou, Elena Anagnostopoulou, and Florian Schäfer
- 56 Control and Restructuring
by Thomas Grano
- 57 The Interaction of Focus, Givenness, and Prosody
A Study of Italian Clause Structure
by Vieri Samek-Lodovici
- 58 The Morphosyntax of Gender
by Ruth Kramer

- 59 The Morphosyntax of Imperatives
by Daniela Isac
- 60 Sentence and Discourse
edited by Jacqueline Guéron
- 61 Optimality-Theoretic Syntax, Semantics, and Pragmatics
From Uni- to Bidirectional Optimization
edited by Géraldine Legendre, Michael T. Putnam, Henriëtte de Swart, and Erin Zaroukian
- 62 The Morphosyntax of Transitions
A Case Study in Latin and Other Languages
by Victor Acedo-Matellán
- 63 Modality Across Syntactic Categories
edited by Ana Arregui, María Luisa Rivero, and Andrés Salanova
- 64 The Verbal Domain
edited by Roberta D'Alessandro, Irene Franco, and Ángel J. Gallego
- 65 Concealed Questions
by Ilaria Frana
- 66 Parts of a Whole
Distributivity as a Bridge Between Aspect and Measurement
by Lucas Champollion

Published in association with the series
The Oxford Handbook of Linguistic Interfaces
edited by Gillian Ramchand and Charles Reiss

In preparation

- Phi Syntax
A Theory of Agreement
by Susana Béjar
- Stratal Optimality Theory
by Ricardo Bermúdez Otero
- Phonology in Phonetics
by Abigail Cohn
- Quantity Superlatives and Proportional Quantification
by Carmen Dobrovie-Sorin and Ion Giurgea
- Lexical Semantics and Morphosyntactic Variation
by Itamar Francez and Andrew Koontz-Garboden
- Generality and Exception
by Iván García-Álvarez
- Computing Optimality
by Jason Riggle
- Pragmatic Aspects of Scalar Modifiers
by Osamu Sawada
- Gradience in Split Intransitivity
by Antonella Sorace
- The Structure of Words at the Interfaces
edited by Heather Newell, Máire Noonan, Glyne Piggott, and Lisa deMena Travis